

# How Much Therapy Is Really Enough?

## *A Session-by-Session Analysis of the Psychotherapy Dose-Effect Relationship*

SCOTT W. KADERA, B.S.

MICHAEL J. LAMBERT, PH.D.

ALISON A. ANDREWS, B.S.

*Previous investigations of the dose-effect relationship in psychotherapy suggest that therapeutic benefits occur early in treatment.*

*Approximately 25% of patients have been estimated to improve after 1 session, and 50% improve by 8 sessions. The aim of this study was to compare such estimates with the actual performance of outpatients in therapy. Forty-five patients seen at an outpatient clinic were monitored session by session for evidence of clinically significant change. Results indicated only 22% of patients "recovered" (as defined in this study) after 8 sessions, with the earliest recovery occurring after 2 sessions. Possible reasons for the lower dose-effect relationship found here are discussed.*

(The Journal of Psychotherapy Practice and Research 1996; 5:132-151)

One of the earliest studies to have sparked researchers' interest in the relationship between treatment duration and therapeutic benefit was that of Seeman.<sup>1</sup> He examined a small sample of college students and concluded that therapeutic gain, based on therapists' assessments, was greatest for patients remaining in therapy for at least 20 sessions. This work was followed by several studies using only slight variations in methodology and yielding similar results.<sup>2-6</sup> These investigations unanimously demonstrated that therapeutic benefit is positively and significantly related to length of psychotherapy, a finding consistent with current reviews of the literature.<sup>7</sup> However, the methodology employed is inadequate by today's standards, and exclusive reliance on patient samples drawn from university counseling centers limits the generalizability of the results.

A contemporary resurgence of interest in the relationship between therapeutic benefit and treatment duration was stimulated by the oft-cited study of Howard et al.<sup>8</sup> These inves-

---

Received March 17, 1995; revised July 11, 1995; accepted July 20, 1995. From the Department of Psychology and Clinical Psychology, Brigham Young University, 284 Taylor Building, P.O. Box 28626, Provo, Utah 84602-8626. Direct correspondence to Dr. Lambert at the above address.

Copyright © 1996 American Psychiatric Press, Inc.

tigators proposed a psychotherapy "dosage" model in which dose was defined as the number of sessions of therapy and treatment effect was defined as the percentage of patients expected to be improved by each session. Dose-effect figures were derived by submitting the findings of 15 outcome studies drawn from the literature to probit analysis and computing mean expected improvement rates for each session. Results indicated that approximately 15% of patients could be expected to have improved prior to the first session (presumably through spontaneous remission), 25% would be improved after 1 session, 50% would be improved after 8 sessions, and about 75% would be improved by 26 sessions. Further analysis indicated that depressed patients responded to the lowest dosages of therapy, followed by those with anxiety neuroses, and borderline-psychotic patients required the highest dosage levels for response to treatment.

Intrigued by the differential responses to therapy of these distinct diagnostic groups, Kopta et al.<sup>9</sup> investigated the response rates of individual psychological symptoms to increasing dosages of therapy. Probit analysis was performed on 10 data sets obtained from patients visiting mental health centers who completed a pretest and at least one test during or after therapy (the Symptom Checklist-90-Revised [SCL-90-R]). The therapeutic dose at which each symptom was judged to have remitted sufficiently was determined by application of the criteria for clinically significant change proposed by Jacobson and Truax.<sup>10</sup> Results were empirically arranged into three categories: 1) acute distress symptoms characterized by high remission rates just prior to therapy and 68% to 95% improvement by 52 sessions (about 1 year of therapy); 2) chronic distress symptoms distinguished by low remission rates just prior to therapy and 60% to 86% improvement by 52 sessions; and 3) characterological symptoms showing mixed remission rates prior to therapy and 30% to 59% improvement by 52 sessions.

Interestingly, symptoms of anxiety and

depression were about evenly distributed between the acute distress and chronic distress categories. The logical expectation, based on the earlier response to therapy by depressed patients than patients with anxiety neurosis in the original study,<sup>8</sup> would have been for depressive symptoms to predominate in the acute distress category and anxiety symptoms to predominate in the chronic distress category. Moreover, for the most commonly endorsed symptoms, the average dosage at which 75% of patients showed clinically significant symptom remission was 58 sessions. This is 32 sessions, or about 7 months of therapy, more than estimated by Howard et al.<sup>8</sup> for similar improvement. Such major discrepancies clearly demand further investigation before conclusive generalizations about the dose-effect relationship in psychotherapy can confidently be made.

A focused critique of Howard et al.<sup>8</sup> seems to have been offered by only one reviewer, Phillips.<sup>11</sup> He briefly described "limitations and atypical aspects" of the database and concluded that concrete generalizations about the relation of length of psychotherapy to outcome would be premature given currently available data. To the present authors, several significant limitations of the database and methodology of Howard et al. readily come to mind: 1) treatments applied were predominantly psychodynamic or interpersonal, with behavioral, cognitive-behavioral, humanistic-existential, and integrative-eclectic modes of therapy severely underrepresented or excluded; 2) outcome criteria were inconsistent regarding aspects of patient functioning to be measured and vague regarding magnitude of change necessary to indicate significant improvement; 3) measurements were taken predominantly before and after therapy, requiring the assumption of a steady linear progression during therapy and neglecting the possibility that significant improvement was achieved before termination; and 4) patients were classified according to diagnosis; this may not be the most appropriate way to place them into homogeneous groups for compara-

tive purposes, especially in light of the finding that symptoms associated with the same diagnosis respond differentially to treatment.<sup>9</sup>

Howard et al.<sup>8</sup> made important contributions by introducing the concept of dose-effect relationship in psychotherapy, illustrating its practical importance, and demonstrating that empirically based calculations of appropriate lengths of therapy can be realized. Despite the merit of this work, however, curiously little research and criticism has subsequently been devoted to the topic.<sup>9,11,12</sup> One explanation may be that the handful of studies performed to date are meta-analytical. They neither establish procedures nor suggest guidelines for designing an empirical investigation of the psychotherapy dose-effect relationship in a laboratory or field setting. Consequently, in the present study we implemented several novel procedures to obtain the necessary information to calculate an accurate dose-effect relationship with patients studied specifically for that purpose.

First, session-by-session assessment of patients supplemented the usual pretreatment and posttreatment testing used in standard outcome studies. Whereas the independent variable in traditional research is the application of a specific type and minimum course of treatment, the independent variable in dose-effect investigations is the application of successive therapeutic sessions. To adequately assess all levels of this variable, we obtained patient data for each session. We conjectured that the mathematical procedure of extrapolating from pretreatment and posttreatment measurements, because it is based on the assumption of a straight-line model of change that ignores patients' unstable, nonlinear responses to therapy,<sup>13</sup> might result in dose-effect figures that misrepresented patients' actual needs for psychotherapy.

Second, the criteria for clinically significant change developed by Jacobson and Truax<sup>10</sup> were employed to make judgments of individual patient outcomes. These require 1) that a patient's scores on a given instrument progress during therapy from the score distri-

bution of dysfunctional peers into the score distribution of functional peers and 2) that such change be of sufficient magnitude to be statistically reliable. Patients meeting both criteria are considered "recovered" for the purpose of analysis. Such a standard is more stringent than that applied in the database of Howard et al.<sup>8</sup> However, clinical-significance criteria were used by Kopta et al.,<sup>9</sup> which may explain the large difference (7 months) between the two studies in the estimated number of sessions required for 75% of patients to benefit from therapy. The use of clinical-significance methodology in the present study dictated that only data from patients initially categorized as dysfunctional could be used in the derivation of dose-effect figures. This conservative approach ensured that the results adequately reflected the therapeutic needs of the more severely ill patients not only to improve, but to return to relatively normal functioning.

Third, in the present investigation we used a nonexperimental design that omitted the usual methodological requirement of a control group. Speer<sup>14</sup> recently questioned the effectiveness of traditional experimental outcome research as a means of informing policymakers and practitioners. Speer demonstrated that single-sample, nonexperimental field investigations may provide more readily understandable findings with increased external validity. This may be the design of choice for dose-effect research, which requires methods deemed superfluous by standard outcome studies (session-by-session assessment) and omits other methods that such studies include. For example, the addition of a control group in a dose-effect study would allow the conclusion that patients recovered as a consequence of therapeutic interventions rather than other causes, but little inferential usefulness is added to the main object of inquiry, the establishment of a dose-effect relationship.

The main purpose of the study was to test the predictive accuracy of previous estimates of the psychotherapy dose-effect relationship<sup>8,9</sup> by comparing them with the therapeutic dosages required for clinically significant change of patients seen in a typical setting. To

this end, session-by-session changes in the severity of symptoms of a small group of patients seen at an outpatient clinic were analyzed to yield dose-effect figures for comparative purposes. A secondary purpose of the study was to explore the usefulness of the three procedures discussed above for conducting research into the nature of the dose-effect relationship in psychotherapy.

## METHODS

### Participants

Patients were 64 adults who received psychotherapy from a university-based outpatient clinic that is used as a training facility for clinical psychology and social work graduate students. The clinic offers psychological services free of charge and serves a county of approximately 265,000 residents. Because services are free, the clinic attracts clients with lower incomes or without mental health insurance benefits. Relatively few college students receive psychotherapy at the clinic because most are referred to the free student services provided by the university's counseling and development center.

A total of 82 adults requested services over a 10-month period. Five refused to participate in the study, leaving an original pool of 77 patients who read and signed an informed consent form regarding their participation in the study. Of these, 13 did not continue to seek therapy after completing the pretest, which consisted of the Outcome Questionnaire (OQ) discussed below. The OQ pretest scores of these study dropouts did not differ significantly from the OQ pretest scores of the study sample ( $t = 0.51, P = 0.6$ ).

Demographic characteristics of the 77 patients in the original pool and the 64 patients in the study sample are presented in Table 1. Patients who participated in the study were primarily female, relatively young, college educated, and of lower income. Among males, marital status was about evenly distributed; among females, the majority were single or

married, with relatively few divorced. A mood disorder was the predominant diagnosis for both males and females. The breakdown of principal diagnoses for the sample was as follows: major depression, 36%; dysthymia, 11%; depressive disorder not otherwise specified, 6%; bipolar disorder, 6%; adjustment disorder, 3%; posttraumatic stress disorder (PTSD), 5%; panic disorder, 3%; simple phobia, 3%; social phobia, 2%; personality disorder, 6%; V code, 13%; and other, 6%.

Therapy was conducted by 23 clinical psychology and 13 social work trainees in accredited graduate programs. Clinical therapists were 6 third-year students (4 male, 2 female), 8 second-year students (5 male, 3 female), and 9 first-year students (7 male, 2 female). Social work therapists were 13 first-year students (6 male, 7 female). Each therapist received a minimum of 1 hour of weekly supervision. Patient diagnoses were given by the case therapist on the basis of a clinical interview and whatever additional testing was deemed necessary by the therapist and supervisor. For patients seen by social work trainees, diagnoses were assigned by the authors from a review of documents in the patient's case file, such as intake interview, social history, and case notes. The diverse theoretical orientations of the supervisors and trainees included cognitive-behavioral, humanistic-existential, psychodynamic-interpersonal, and integrative-eclectic approaches. The majority of patients (78%) were seen by the 23 therapists in the clinical psychology program.

Of some concern was whether student therapists would be able to provide the same quality of intervention as professional therapists. A recent review of the literature<sup>15</sup> noted that the effects on patient outcome of such variables as level of professional training, amount of experience, and professional discipline were equivocal and contradictory. Skill level achieved by the therapist was suggested as possibly the most decisive personal variable in effecting change, but the relationship between amount of training and acquired skill remained uncertain. Speer<sup>14</sup> detected no sig-

nificant difference in outcomes between student therapists and more experienced professionals, and a major recent meta-analytical review<sup>16</sup> found only "modest" effect sizes for the efficacy of more trained over less trained therapists. Still, the use of trainees in this study is an admitted shortcoming in design, and its ultimate effect on the generalizability of the results is uncertain.

#### Assessment

The Outcome Questionnaire recently developed by Lambert et al.<sup>17</sup> was used to assess

patient change. The questionnaire contains 45 items (five-point scale) that measure patient functioning in the areas of major psychiatric symptoms, social-role functioning, and interpersonal relationships. Symptom distress items are heavily oriented toward symptoms of anxiety and depression, and they also assess substance abuse, stress, and more severe pathology. Social role items assess work relations and leisure. Interpersonal relationship items assess problems in and satisfaction with friendships, marriage, and family life. It has been suggested that such multidimensional assessment of patient functioning is necessary to

**TABLE 1. Demographic characteristics of patients who sought psychotherapy**

	Total Seeking Therapy (n = 77)		Total Included in Study (n = 64)		Dysfunctional Group (n = 45)		Functional Group (n = 19)	
	M	F	M	F	M	F	M	F
Gender (%)	31	69	31	69	31	69	42	58
Age (years, mean ± SD)	32 ± 10	28 ± 9	33 ± 10	28 ± 9	33 ± 11	28 ± 9	32 ± 9	29 ± 9
Race (%)								
Caucasian	100	94	100	95	100	93	100	100
Hispanic	0	4	0	5	0	7	0	0
Asian	0	2	0	0	0	0	0	0
Marital Status (%)								
Single	35	55	37	55	33	58	43	44
Married	39	33	37	36	33	33	43	44
Divorced/separated	26	12	26	10	33	9	14	11
Education (%)								
0-12 yrs	26	20	21	14	33	15	0	11
13-16 yrs	61	79	69	83	58	82	86	89
17+ yrs	13	2	11	2	8	3	14	0
Income (%)								
< \$10,000/yr	40	38	36	41	30	39	43	44
\$10,000-\$20,000	30	26	29	28	30	32	29	11
\$20,000-\$30,000	30	27	36	27	40	21	28	44
> \$30,000	0	8	0	6	0	8	0	0
Diagnosis (%)								
Mood disorder	NA	NA	55	63	75	73	25	36
Anxiety disorder	NA	NA	0	14	0	15	0	9
Personality disorder	NA	NA	10	5	8	6	12	0
Other	NA	NA	35	18	17	6	63	55

◆ Note: M = male; F = female; NA = not applicable.

capture the various stages in the apparent stepwise change process in psychotherapy.<sup>18</sup> The total score on the OQ can range from 0 to 180; the severity of reported psychopathology corresponds with increases in the total score.

Psychometric testing of the OQ yielded an internal consistency of 0.93, 3-week test-retest reliability of 0.84, and concurrent validities with similar instruments (such as the Zung Self-Rating Depression Scale, Taylor Manifest Anxiety, SCL-90-R, Social Adjustment Scale, Inventory of Interpersonal Problems) in the range of 0.53 to 0.88. Normative data from a community sample ( $n = 102$ ) produced a mean of  $48.16 \pm 18.23$  (SD). Normative data from a mental health center ( $n = 100$ ) produced a mean of  $86.07 \pm 19.33$ . These samples were defined in the present study as the functional and dysfunctional distributions, respectively, for making estimations of patient change, thus avoiding difficulties in interpretation that arise from shifting cutoff scores when study-specific dysfunctional distributions are used.<sup>10</sup> Patients in the study sample were similar in age, gender, and socioeconomic status to subjects in the normative community and mental health center samples.

#### Procedures

Patients completed the OQ prior to each weekly therapy session. Completion of the pretest occurred immediately before the first session; the first posttest then preceded the second session, the second posttest preceded the third session, and so on. This procedure was consistent with OQ instructions asking patients to describe their functioning "over the last week." Patients received an OQ from the clinic receptionist at the time of their appointment, completed it in a waiting area, and returned it to the receptionist before beginning their session. Patients forgetting to complete an OQ were immediately sent one in the mail and asked to complete it according to their best estimate of how they were functioning prior to their last session. Approximately 3.5%

of obtained questionnaires were completed late in this way.

Decisions to terminate therapy were made at the discretion of the patients and their therapists. At no time during the study was information from the OQs shared with the therapist or the patient; thus, the decision to terminate was not due to feedback from the OQ.

Dosage in this study was measured in terms of the number of sessions received by a patient, one session being defined as the unit of treatment. This method is consistent with previous studies of the dose-effect relationship<sup>8,9</sup> and is the most convenient way to compare dosages across various therapeutic modalities. It contains the assumption that the effective components of a particular therapy are randomly distributed among the number of sessions received and that a patient's total exposure to these effective components increases with each successive session. Effect or response was defined, again following the lead of previous studies, as the cumulative percentage of patients judged to be recovered at each dose or session.

Judgment of patient outcome was made according to methods developed by Jacobson and Truax<sup>10</sup> for making determinations of individual clinically significant change. These included 1) movement of a patient's total score on the OQ from the distribution of dysfunctional peers (mental health center sample) into the distribution of total scores for functional peers (community sample) and 2) a minimum magnitude of change specified by a reliable change index (RCI) to ensure that such change is statistically reliable. Because the dysfunctional distribution (mean  $\pm$  SD =  $86.07 \pm 19.33$ ) and the functional distribution ( $48.16 \pm 18.23$ ) defined in this study possess equal variances, the point of separation between the two distributions, or cutoff score, was calculated by taking the mean of the distribution means.<sup>10</sup> This yielded a cutoff score of 67 or greater for inclusion in the dysfunctional distribution on the OQ. Of the 64 patients in the study, 45 had OQ pretest scores in the dysfunctional range

and 19 were within the functional range. Demographic characteristics of these two groups are presented in Table 1.

The formulas used in the calculation of the RCI appear in Jacobson and Truax.<sup>10</sup> A slight modification was made in the computation of the standard error of difference ( $S_{diff}$ ) used in the denominator of the formula. Because the OQ measures changes over an interval as short as 1 week but was test-retested over 3 weeks, the reported reliability estimate of 0.84 is probably lower than the true reliability, which would have been more closely approximated by a shorter test-retest interval. Tingey et al.<sup>19</sup> recommend use of the internal-consistency reliability estimate when appropriate test-retest figures are not available. Therefore, an alpha of 0.93 was used in the computation of the standard error of difference. The resulting RCI indicated that movement of at least 15 points on the OQ was necessary for patient change to be considered statistically reliable.

To summarize, patients were considered "recovered" when they met both of the criteria for clinically significant change by 1) moving from the OQ dysfunctional distribution into the OQ functional distribution and 2) showing positive gains of sufficient magnitude to be considered statistically reliable (improvement of at least 15 OQ points). However, since the aim of this study was not only to assess whether a patient had recovered, but also to indicate when that recovery occurred in order to compute a dose-effect relationship, a third criterion had to be specified. Session-by-session assessment of change raised the possibility that some patients might be observed continuing in therapy after obtaining recovered status or might fluctuate between recovered and unrecovered status prior to termination. Therefore, patients were considered recovered at the earliest session at which they persistently met the criteria for clinically significant change. Patients who at some point in therapy met the criteria for clinically significant change but who eventually terminated in the dysfunctional range or failed to improve by at least 15 OQ points were considered not

recovered in the calculation of dose-effect figures.

## RESULTS

OQs were obtained for approximately 97% of the 805 sessions conducted during the study. Therapy outcome status for the 64 patients is summarized in Table 2. "Recovered" patients, as discussed, met both criteria for clinically significant change. "Improved" patients met the criterion for statistical reliability by improving by at least 15 OQ points but remained within the same dysfunctional or functional distribution they were in before starting therapy. "Deteriorated" patients moved at least 15 OQ points in the direction of increasing psychopathology. Patients showing "no change" did not improve or deteriorate more than 15 OQ points during therapy. Results indicate 21 patients (33%) recovered, 16 patients (25%) improved, 24 patients (37%) experienced no change, and 3 patients (5%) deteriorated. Thus, about half (58%) of the patients showed reliably positive gains during therapy.

The percentages for improvement (58%) and deterioration (5%) obtained in this study are fairly consistent with those obtained in other studies with outpatients.<sup>14,20,21</sup> This finding mitigates the concern that the student therapists may have been deficient in the application of effective interventions compared with more experienced professionals.

The negative effect of psychotherapy on some patients is a repeated finding in previous research; less deterioration usually occurs in control groups than in treated samples.<sup>22</sup> However, it is often difficult to distinguish those negatively affected by psychotherapy from those already on a progressive decline and immune to therapeutic interventions. The 3 patients (5%) who deteriorated were all diagnosed with mood disorders and all began therapy within the dysfunctional distribution.

The 21 recovered patients represent 47% of those categorized as dysfunctional at the beginning of treatment. These patients were used to formulate an initial dose-effect rela-

TABLE 2. Outcome status for 64 patients receiving psychotherapy

Patient	Outcome Questionnaire Scores			Outcome
	Pretest	Posttest	Change	
1	95	57	38	Recovered
2	85	70	15	Improved
4	96	70	26	Improved
5	51	31	20	Improved
7	64	29	35	Improved
8	41	46	-5	No change
9	41	51	-10	No change
10	82	47	35	Recovered
11	83	32	51	Recovered
12	115	50	65	Recovered
13	126	59	67	Recovered
15	29	31	-2	No change
17	74	61	13	No change
18	73	53	20	Recovered
19	105	72	33	Improved
20	30	38	-8	No change
21	108	127	-19	Deteriorated
22	88	65	23	Recovered
23	75	41	34	Recovered
24	89	28	61	Recovered
25	62	22	40	Improved
26	78	100	-22	Deteriorated
27	102	50	52	Recovered
28	70	66	4	No change
29	90	64	26	Recovered
30	116	49	67	Recovered
31	56	45	11	No change
33	124	56	68	Recovered
34	93	76	17	Improved
35	32	5	27	Improved
36	94	90	4	No change
37	96	48	48	Recovered
38	69	52	17	Recovered
40	107	93	14	No change
41	79	46	33	Recovered
42	88	73	15	Improved
43	33	12	21	Improved
44	78	41	37	Recovered
45	97	76	21	Improved
46	91	72	19	Improved
48	66	42	24	Improved
49	112	114	-2	No change
50	63	38	25	Improved
51	85	60	25	Recovered
52	60	51	9	No change
53	118	98	20	Improved
54	87	31	56	Recovered
55	84	58	26	Recovered
56	58	58	0	No change
57	75	61	14	No change
59	105	95	10	No change

*(continued)*



TABLE 2. Outcome status for 64 patients receiving psychotherapy (continued)

Patient	Outcome Questionnaire Scores			Outcome
	Pretest	Posttest	Change	
60	90	91	-1	No change
61	85	57	28	Recovered
66	114	140	-26	Deteriorated
68	101	89	12	No change
71	83	78	5	No change
72	69	59	10	No change
73	87	73	14	No change
74	81	80	1	No change
75	17	3	14	No change
76	37	23	14	No change
77	53	41	12	No change
78	35	34	1	No change
79	52	37	15	Improved

tionship, which is graphically displayed in Figure 1. The minimum number of sessions required for patient recovery was 2, and all patients who recovered did so by 25 sessions. Approximately 14% of recovering patients were recovered by 4 sessions, 43% by 8 sessions, and 76% by 13 sessions. This curve should not be interpreted as reflecting the typical occurrence of recovery for the dysfunctional patient sample as a whole because it was constructed solely from those who recovered. Rather, the curve represents the time frame of recovery for the 21 dysfunctional patients who did recover.

The Howard et al.<sup>8</sup> dose-effect analysis assumes a hypothetical population of patients remaining in therapy until improved. To calculate a comparable dose-effect relationship in this study, a second type of analysis was performed that considered all 45 patients in the dysfunctional sample. Outcome data were arranged in the form of a recovery table similar to mortality tables constructed by insurance companies to predict death rates. These data are presented in Table 3. One limitation in adopting a method used to construct mortality rates is that whereas everyone will eventually die, not all those receiving psychotherapy will eventually change. Nevertheless, the table provides a useful way of analyzing session-by-session outcome data with patients completing different lengths of therapy.

In Table 3, the first four columns are based on the observed data, and the next three columns are hypothetical constructions. The column labeled "Patients Treated" indicates the number of patients who received  $x$  sessions of therapy, regardless of their eventual outcome. The column labeled "Recovered" indicates the number of patients who recovered following  $x$  sessions.  $P_x$  represents the proportion of patients recovered following session  $x$  calculated from the previous two columns. For example, after the fifth session, 2 patients met the criteria for recovery out of a total of 38 patients receiving 5 sessions. This results in a recovery rate ( $P_x = 2/38 = 0.053$ ) of approximately 5% following session 5.

Figures from column  $P_x$  were then used to construct the data for columns  $T_x$  and  $R_x$ .  $T_x$  represents the number of patients in a hypothetical treatment population of 1,000 who have not yet recovered.  $R_x$  represents the number of patients from the same population who would be expected to recover based on the observed proportion recovered from column  $P_x$ . For example, of 1,000 patients receiving 2 therapy sessions, 23 patients are expected to recover based on the  $P_x$  of 0.023, leaving 977 patients to continue with a third session. Of these, 24 ( $977 \times 0.024$ ) are expected to recover following the third session, leaving 954 to continue with a fourth session, and so on.

Column  $Ex$  represents the expected additional number of sessions needed for recovery by a patient already receiving  $x$  sessions. This figure was derived by using the weighted average number of sessions required for recovery at each session. For example, according to column  $Ex$ , a patient just beginning therapy can be expected to need about 11 sessions, and a patient completing 1 session but not yet recovered will probably need about 10 additional sessions.

By cumulatively summing column  $Rx$  at each session and dividing by the original treatment population ( $Tx$ ) of 1,000, we arrived at a dose-effect recovery curve that is comparable to the dose-effect figures of Howard et al.<sup>8</sup> This curve represents the percentage of patients expected to be recovered at each session from a hypothetical treatment population continuing indefinitely in therapy, based on the proportion of patients recovered at each session in the present study. The resulting curve is

**TABLE 3. Recovery table based on psychotherapy performance of 45 dysfunctional patients**

Session #	Patients Treated	Patients Recovered	$Px^a$	$Tx^b$	$Rx^c$	$Ex^d$
1	45	0	0.000	1,000	0	10
2	44	1	0.023	1,000	23	9
3	42	1	0.024	977	23	8
4	40	1	0.025	954	24	7
5	38	2	0.053	930	49	7
6	34	3	0.088	881	78	7
7	31	0	0.000	803	0	7
8	31	1	0.032	803	26	6
9	26	0	0.000	778	0	6
10	25	2	0.080	778	62	5
11	23	3	0.130	715	93	5
12	20	1	0.050	622	31	6
13	18	1	0.056	591	33	6
14	16	0	0.000	558	0	6
15	15	0	0.000	558	0	5
16	11	1	0.091	558	51	4
17	10	0	0.000	507	0	4
18	9	1	0.111	507	56	3
19	8	0	0.000	451	0	3
20	8	2	0.250	451	113	2
21	4	0	0.000	338	0	4
22	4	0	0.000	338	0	3
23	3	0	0.000	338	0	2
24	3	0	0.000	338	0	1
25	3	1	0.333	338	113	0
26	2	0	0.000	225	0	0

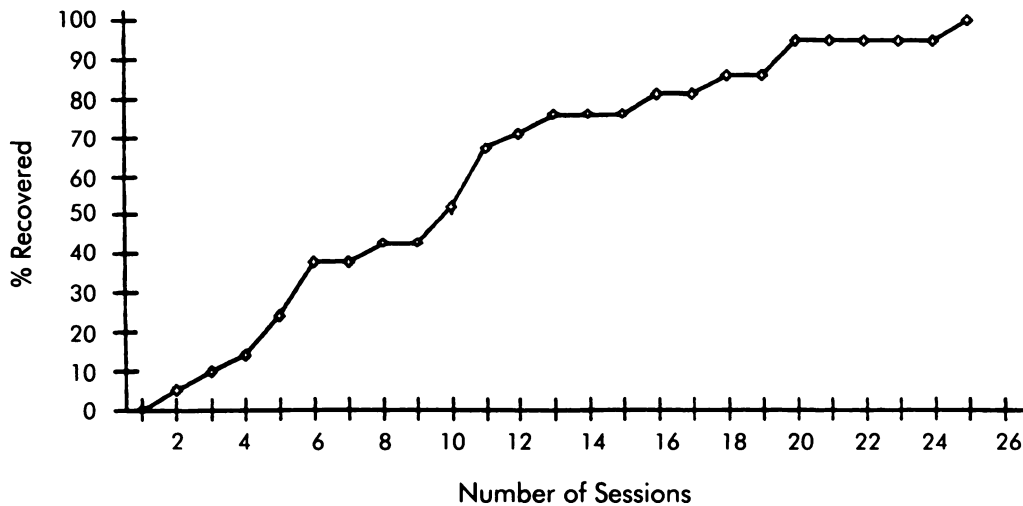
<sup>a</sup> $Px$  represents the probability of patient recovery during session  $x$  based on the observed data (Patients Recovered + Patients Treated).

<sup>b</sup> $Tx$  represents the expected number of unrecovered patients receiving  $x$  sessions from an original hypothetical treatment population of 1,000 patients.

<sup>c</sup> $Rx$  represents the expected number of recovered patients from the hypothetical treatment population ( $Tx$ ) based on the probability of recovery ( $Px$ ).

<sup>d</sup> $Ex$  represents the expected additional number of sessions needed for recovery after receiving  $x$  sessions.

FIGURE 1. Relation of percentage recovered to number of sessions received for 21 previously dysfunctional patients receiving psychotherapy.



presented in Figure 2. Because the curves in Figures 1 and 2 are based on the same outcome data, the curves are similar. However, the inclusion of all 45 dysfunctional patients in the derivation of the latter curve lowers the dose-effect figures considerably. The number of patients expected to be recovered at 4 sessions is 7%; at 8 sessions, 22%; and at 13 sessions, 44%. For comparative purposes, the mean dose-effect figures obtained by Howard et al. from their probit analysis of 15 data sets have also been plotted in Figure 2. The two curves do not begin to converge until about session 20. They are in agreement that by session 26, about 75% of patients can be expected to have significantly recovered.

The therapeutic courses of the 21 recovered patients were plotted and are presented in Figure 3. The horizontal line in each graph represents the minimum OQ score patients needed to obtain, based on their initial OQ pretest score, to meet the criteria for clinically significant change. The shapes of the individual curves are consistent with the previous finding that patients not only show great variability from one another in their responses to therapy, but also show wide fluctuation in their

subjective estimates of the intensity of their symptoms.<sup>13</sup> For example, Patient 1 moved in and out of the functional distribution seven times before meeting criteria for recovery at session 25. Patient 30 reliably deteriorated by session 2, then improved dramatically over the next two sessions to meet criteria for recovery by session 4. Patient 11 made steady improvement after an initial deterioration at session 2, met the criteria for recovery at session 13, but continued in therapy for 16 more sessions with only minor additional benefit. The modal number of sessions a patient remained in therapy after meeting criteria for recovery was 3, with the majority continuing within a range of 0 to 5 sessions.

Examination of the graphs indicates few patients changed in a steady, linear fashion. One reason for employing session-by-session assessment in this study was to accurately determine the session at which patients displaying such individualistic responses to therapy met criteria for clinically significant change. It was conjectured that the method of estimating recovery by extrapolating from pretreatment and posttreatment measurements, on the basis of a straight-line model of change, could result

in dose-effect figures that misrepresented patients' actual therapy session requirements.

To test this hypothesis, outcomes of the 21 recovered patients were reanalyzed by extrapolating as if only pretreatment and post-treatment measurements were known. The linearly estimated session for recovery was then compared with the actual session at which each patient recovered. Results indicated underestimation for 12 patients (median underestimation = 3.5 sessions, range of underestimation = 1-7 sessions), overestimation for 6 patients (median overestimation = 1.5 sessions, range of overestimation = 1-3 sessions), and agreement for 3 patients. The net effect of extrapolating from pre-post measures was to underestimate the dosage requirements of the 21 recovered patients by a mean of 2 sessions. The consequence of pre-post extrapolation in this study, although minor, would have been to inflate the obtained dose-effect figures in the direction of the relationship obtained by Howard et al.<sup>8</sup>

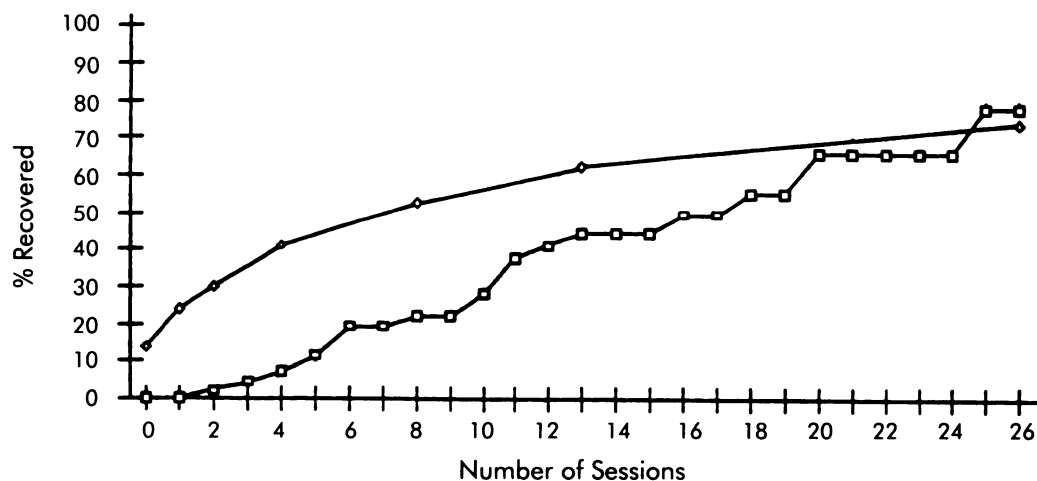
Data from patients whose OQ pretest scores were in the functional range were not used in the derivation of dose-effect figures because these scores indicated a lack of signifi-

cant initial psychopathology. However, 8 of 19 functional patients (42%) "improved" during therapy by exhibiting positive change of at least 15 OQ points, thus highlighting one of the difficulties of placing patients into functional and dysfunctional categories solely on the basis of pretest scores. It would certainly be useful to identify beforehand those persons who do not really need psychotherapy, but, as the present results show, a more precise method is called for than labeling patients as functional on the basis of a single instrument. Functional improvers required a range of 2-13 sessions and mean of 7 sessions for improvement, compared with the range of 2-25 sessions and mean of 11 sessions required for dysfunctional recoverers. This finding is consistent with the notion that less severely ill patients require less therapy for change.

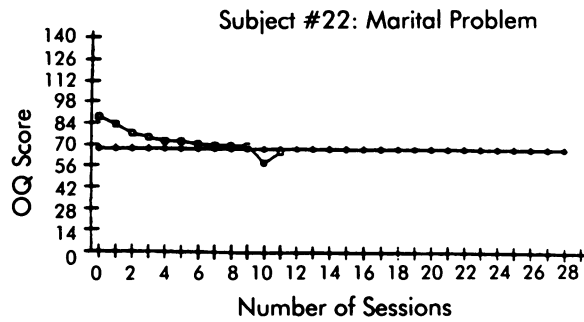
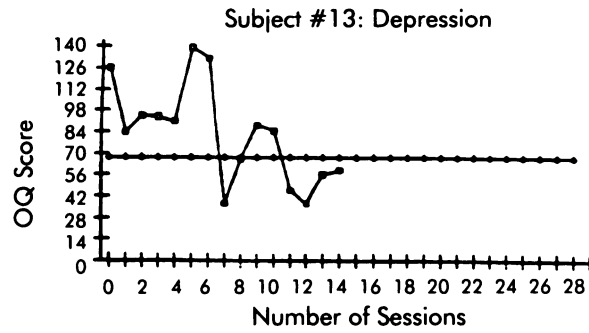
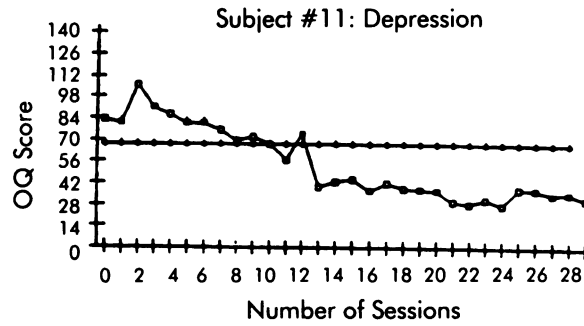
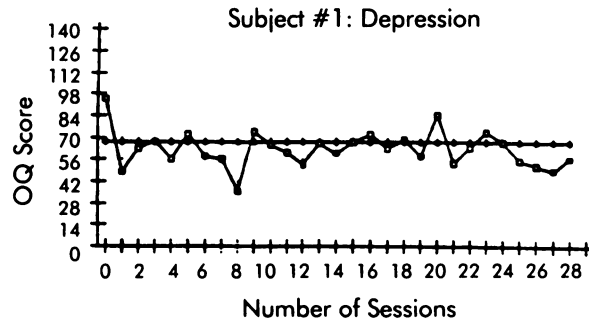
#### DISCUSSION

The results of the present study raise some intriguing questions. Are previous estimates of the dose-effect relationship in psychotherapy<sup>8,9</sup> too optimistic? How appropriate are the

FIGURE 2. Dose-effect estimates calculated from 45 dysfunctional patients receiving psychotherapy (squares) compared with dose-effect estimates of Howard et al.<sup>8</sup> (diamonds).



**FIGURE 3. Plotted psychotherapy courses for 21 recovered patients as measured by the Outcome**



---

Questionnaire (OQ). Diamonds: clinically significant change. Squares: OQ Score. (continued)

---

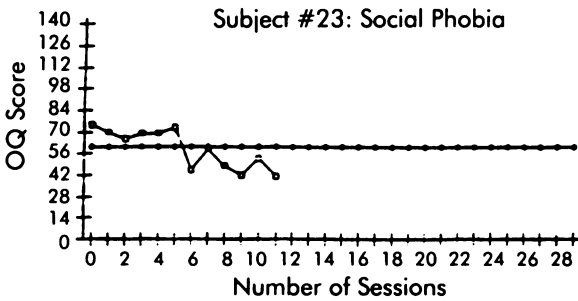
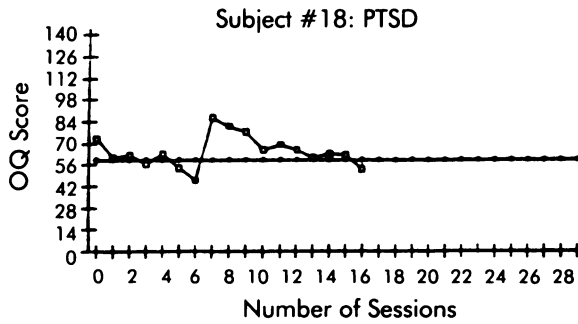
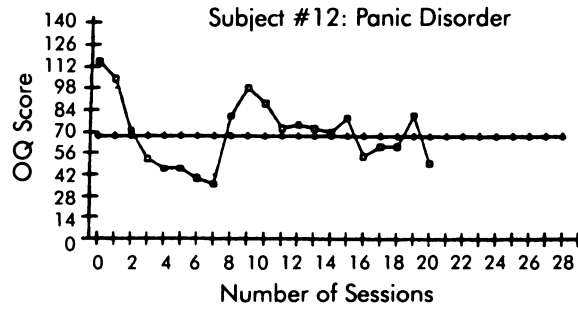
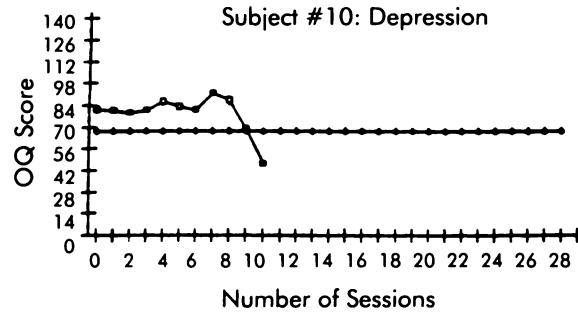
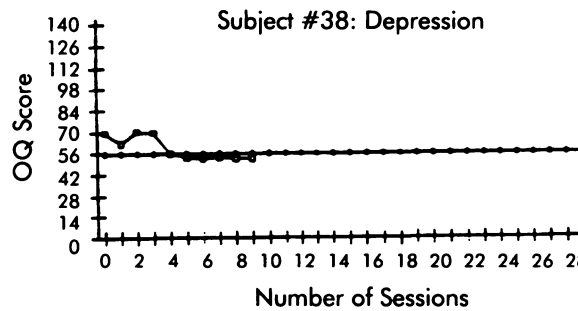
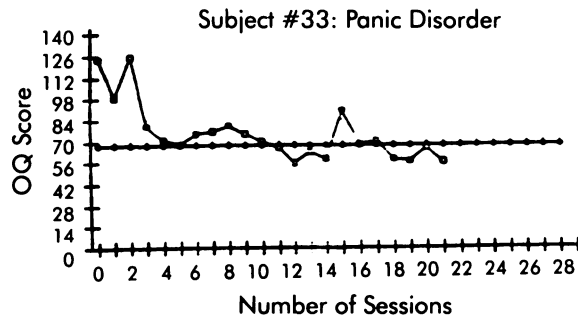
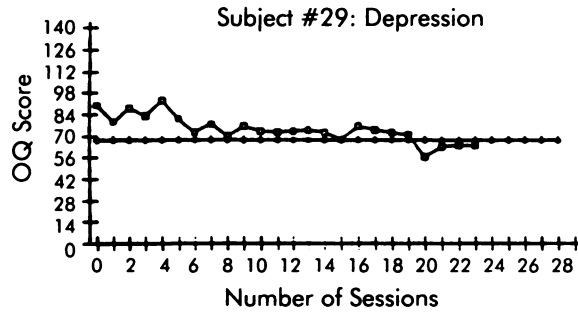
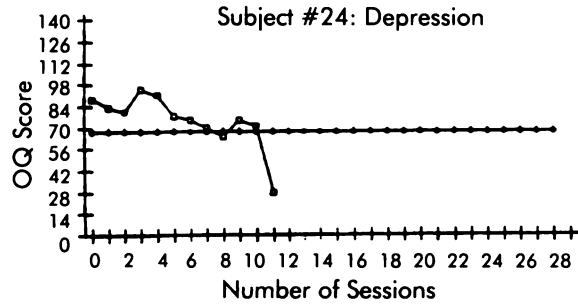


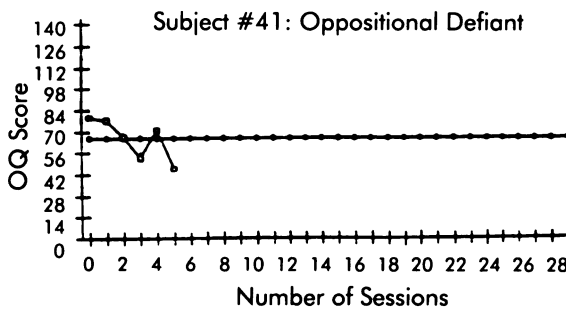
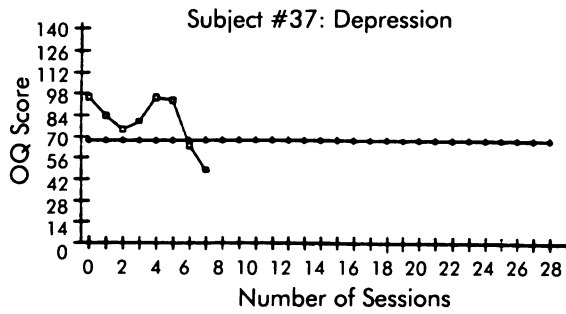
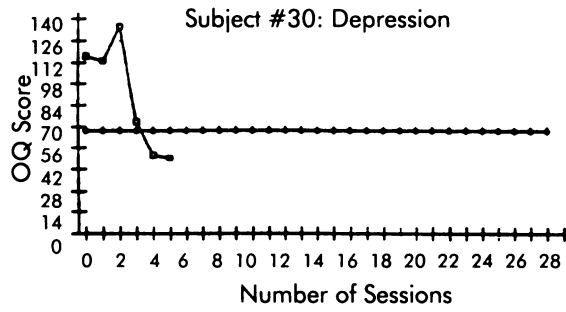
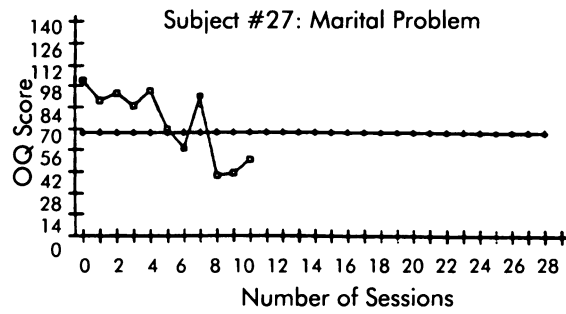
FIGURE 3. Plotted psychotherapy courses for 21 recovered patients as measured by the Outcome



---

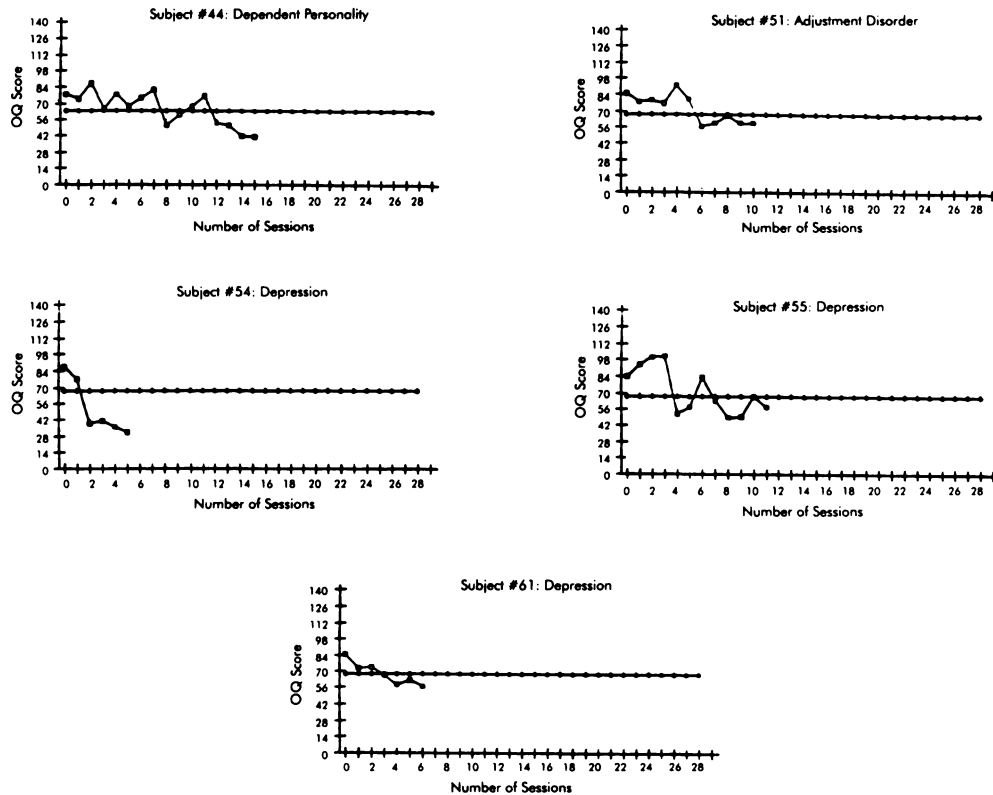
Questionnaire (OQ). Diamonds: clinically significant change. Squares: OQ Score. (continued)

---





**FIGURE 3. Plotted psychotherapy courses for 21 recovered patients as measured by the Outcome Questionnaire (OQ). Diamonds: clinically significant change. Squares: OQ Score. (continued)**



use of mathematical techniques that estimate therapeutic change across sessions for arriving at accurate dose-effect figures? What types of analysis should be used in making predictions about the expected number of sessions patients will need to recover?

The dose-effect figures derived from the 45 dysfunctional patients (Figure 2) differ significantly from the estimates of previous studies. Howard et al.<sup>8</sup> predict 50% of patients will be measurably improved by 8 sessions, whereas performance of patients in the present study suggests only 22% recovery for the same number of sessions. A 50% rate of recovery is not expected until session 16, twice the number of sessions indicated by Howard et al. Additionally, whereas Howard et al. suggest

15% of patients will be spontaneously improved prior to beginning therapy and Kopta et al.<sup>9</sup> predict spontaneous remission rates in the range of 9% to 50% for particular symptoms, no patient in the present study recovered following session 1, and only 1 patient was recovered by session 2.

The dosage curves in Figure 2 indicate that by session 26, about 75% of patients can be expected to meet the criteria for a successful therapeutic outcome. Unfortunately, none of the patients in this study received more than 29 sessions, so comparisons cannot be made for larger dosages of therapy. Although Kopta et al. reported 75% expected improvement by 58 sessions for the most commonly endorsed symptoms on the SCL-90-R, results of this

study indicate that such a figure may overestimate patients' actual needs. Individual assessment of patients' global scores on the SCL-90-R might have produced a considerably lower dosage requirement for 75% patient improvement than the analysis of only the most commonly endorsed symptoms.

Several reasons can be given to explain the initially lower dose-effect findings obtained here. First, recovery rates in our study were based on the performance of patients initially categorized as dysfunctional. This was an admittedly conservative move that was deemed necessary to ensure the dose-effect figures adequately reflected the therapeutic needs of the more severely ill patients in the study. Those comprising the Howard et al.<sup>8</sup> database, if classified according to currently employed clinical significance criteria, would probably represent a mixture of both dysfunctional and functional patients. Functional patients in this study required a mean of 7 sessions for improvement and were improved by 13 sessions. In contrast, dysfunctional patients required a mean of 11 sessions and used a maximum of 25 sessions. Inclusion of functional patients in the analysis would have shifted the dose-effect curve in the direction of the relationship obtained by Howard et al., thus indicating earlier initial responses to therapy.

Second, in our study therapeutic progress was assessed with a reliable, valid, standardized outcome instrument.<sup>17</sup> Patient functioning was measured in the areas of symptom distress, social-role functioning, and interpersonal relationships, and strict standards were applied for judging clinically significant change.<sup>10</sup> In the Howard et al.<sup>8</sup> database, no consistent criteria regarding aspects of patient functioning to be measured or magnitude of change necessary to indicate a significant response to therapy seem to have been used in making decisions about patient improvement. Reanalysis of these patients might reveal that some judged as improved fell short of the stricter criteria of recovery employed in this study, thus lowering the estimated dose-effect relationship.

Third, dose-effect figures were calculated from data of patients assessed session by session rather than from mathematical extrapolations across sessions. The danger in making assumptions about therapeutic movement between pretreatment and posttreatment measurements is dramatically emphasized by examining the therapeutic course of Patient 24, shown in Figure 3. Clinically significant change was obtained at session 11. However, the assumption of steady linear improvement from session 0 to session 11 would indicate recovery occurred at session 4, or about two-thirds earlier in the treatment than when the patient did recover. Reanalysis of the 21 recovered patients in this study using mathematical extrapolation from pretreatment and post-treatment data resulted in an underestimation of the actual therapy needs for 12 patients. The median underestimation for these 12 patients was 3.5 sessions. The median underestimation when calculated for the 21 patients as a whole was about 2 sessions. A similar phenomenon may have been operating in the Howard et al.<sup>8</sup> and Kopta et al.<sup>9</sup> analyses, making patients and psychological symptoms, respectively, appear to have improved sooner than they actually did.

Although the dose-effect figures obtained here are lower for earlier sessions than those of previous studies,<sup>8,9</sup> several methodological factors should be mentioned that probably tended to inflate or accelerate the recovery figures actually reported in this study. These included 1) use of the internal consistency reliability in the computation of the reliable change index in place of the lower test-retest reliability estimate (lower estimates of reliability reduce the percentage of patients judged improved); 2) classification of patients as recovered at the earliest session when continuous clinically significant change was first realized rather than at the time of termination; and 3) accelerated decrease of the denominator used in the calculation of  $P_x$  in the recovery table (Table 3) caused by the withdrawal from therapy of nonrecovered patients. Alteration of these procedures would not only tend to

decrease the slope of the recovery curve, but would also diminish the number of patients judged to have recovered.

By and large, the methodology employed in this study (strict criteria for recovery, session-by-session assessment, and single-sample design) seems to have been useful in obtaining and analyzing the information necessary to calculate dose-effect figures for patients studied for that purpose in a natural setting. We hope that the accumulation of similar data from future studies will yield a database from which estimates of actual patient needs for therapy may accurately and confidently be made. Once larger numbers of patients have been studied with this methodology, several aspects of the foregoing analyses may prove useful to third-party payers in establishing equitable guidelines for financial support of psychotherapy or may be of assistance to practicing clinicians in formulating adequate treatment plans.

For example, the recovery table shown in Table 3 can be used to assess the likelihood of patient recovery at each session as well as to estimate the additional number of sessions needed for those not yet recovered. The type of curve presented in Figure 3 can provide general dose-effect recovery estimates from among all those entering psychotherapy and is similar to the original dosage curve provided by Howard et al.<sup>8</sup> The type of curve presented in Figure 1 affords an expected time frame for recovery based on the therapeutic performance of previous recoverers. It is not hard to imagine various recovery curves of this latter sort for distinct patient groups that could be used both to determine optimal dosages of therapy and to identify potential nonrecoverers.

The findings of this study are consistent with past research in demonstrating that empirically based dosage estimates can be determined and will prove useful for clinical practice and public policy. These findings also highlight the need for more studies devoted explicitly to an investigation of the dose-effect relationship in psychotherapy. This study appears to be the first of its kind to directly examine patients for the sole purpose of assessing the dose-effect relationship. Future research should be conducted in other typical treatment settings with more experienced professionals and a larger number of patients.

Future research should also include a variety of measures of outcome rather than a single self-report scale and should combine the views of patients and observers to effectively assess a wider range of mental health components as well as capture the different stages of change in the therapeutic process.<sup>23</sup> This research should consider the most meaningful way to classify patients for comparative purposes. It would be important to know whether different patient subtypes require different dosages of therapy. Do some disorders require more treatment on average to effect change than others? How much more? It would also be important to know whether therapies intended to be brief are any more efficient than interventions having no theoretical or practical time constraints. What are the most efficient therapeutic interventions and modalities?

By indicating how little is known about this still-emerging topic in psychotherapy outcome research, such questions affirm the importance of continuing study into the dose-effect relationship.

---

#### R E F E R E N C E S

1. Seeman J: Counselor judgments of therapeutic process and outcome, in *Psychotherapy and Personality Change*, edited by Rogers CR, Dymond RF. Chicago, University of Chicago Press, 1954, pp 99-108
2. Cartwright DS: Success in psychotherapy as a function of certain actuarial variables. *J Consult Psychol* 1955; 19:357-363
3. Standal SW, van der Veen F: Length of therapy in relation to counselor estimates of personal integration and other case variables. *J Consult Psychol* 1957; 21:1-9
4. Johnson RW: Number of interviews, diagnosis, and

- success of counseling. *Journal of Counseling Psychology* 1965; 12:248-251
5. Weitz LJ, Abramowitz SI, Steger JA, et al: Number of sessions and client-judged outcome: the more the better? *Psychotherapy: Theory, Research and Practice* 1975; 12:337-340
  6. Strassberg DS, Anchor KN, Cunningham J, et al: Successful outcome and number of sessions: when do counselors think enough is enough? *Journal of Counseling Psychology* 1977; 24:477-480
  7. Orlinsky DE, Grawe K, Parks BK: Process and outcome in psychotherapy: *noch einmal*, in *Handbook of Psychotherapy and Behavior Change*, 4th edition, edited by Bergin AE, Garfield SL. New York, Wiley, 1994, pp 270-376
  8. Howard KI, Kopta SM, Krause MS, et al: The dose-effect relationship in psychotherapy. *Am Psychol* 1986; 41:159-164
  9. Kopta SM, Howard KI, Lowry JL, et al: Patterns of symptomatic recovery in psychotherapy. *J Consult Clin Psychol* 1994; 62:1009-1016
  10. Jacobson NS, Truax P: Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; 59:12-19
  11. Phillips EL: Length of psychotherapy and outcome: observations stimulated by Howard, Kopta, Krause and Orlinsky. *Am Psychol* 1988; 43: 669-670
  12. McNeilly CL, Howard KI: The effects of psychotherapy: a reevaluation based on dosage. *Psychotherapy Research* 1991; 1:74-78
  13. Barkham M, Stiles WB, Shapiro DA: The shape of change in psychotherapy: longitudinal assessment of personal problems. *J Consult Clin Psychol* 1993; 61:667-677
  14. Speer DC: Can treatment research inform decision makers? Nonexperimental method issues and examples among older outpatients. *J Consult Clin Psychol* 1994; 62:560-568
  15. Beutler LE, Machado PP, Neufeldt SA: Therapist variables, in *Handbook of Psychotherapy and Behavior Change*, 4th edition, edited by Bergin AE, Garfield SL. New York, Wiley, 1994, pp 229-269
  16. Stein DM, Lambert MJ: Graduate training in psychotherapy: are therapy outcomes enhanced? *J Consult Clin Psychol* 1995; 63:182-196
  17. Lambert MJ, Lunnen K, Umphress V, et al: Administration and Scoring Manual for the Outcome Questionnaire (OQ-45.1). Salt Lake City, UT, IHC Center for Behavioral Healthcare Efficacy, 1994
  18. Howard KI, Lueger RJ, Maling MS, et al: A phase model of psychotherapy outcome: causal mediation of change. *J Consult Clin Psychol* 1993; 61:678-685
  19. Tingey R, Lambert MJ, Burlingame GM, et al: Assessing clinical significance: proposed extensions to method. *Psychotherapy Research* (in press)
  20. Jacobson NS, Wilson L, Tupper C: The clinical significance of treatment gains resulting from exposure-based interventions for agoraphobia: a reanalysis of outcome data. *Behavior Therapy* 1988; 19:539-552
  21. Jacobson NS, Follette WC, Revenstorf D, et al: Variability in outcome and clinical significance of behavioral marital therapy: a reanalysis of outcome data. *J Consult Clin Psychol* 1984; 52:497-504
  22. Lambert MJ, Bergin AE: The effectiveness of psychotherapy, in *Handbook of Psychotherapy and Behavior Change*, 4th edition, edited by Bergin AE, Garfield SL. New York, Wiley, 1994, pp 143-189
  23. Lambert MJ, Hill CE: Assessing psychotherapy outcomes and processes, in *Handbook of Psychotherapy and Behavior Change*, 4th edition, edited by Bergin AE, Garfield SL. New York, Wiley, 1994, pp 72-113