The dose response relationship in psychotherapy: Implications for social policy

PAUL HARNETT¹, ANALISE O'DONOVAN², & MICHAEL J. LAMBERT³

¹School of Psychology, University of Queensland, Brisbane, ²School of Psychology, Griffith University, Brisbane, Australia, and ³Department of Psychology, Brigham Young University, Provo, Utah, USA

Abstract

The principle aim of this study was to estimate the number of sessions of psychotherapy needed for clients suffering from psychiatric illness to return to a normal state of functioning or reliably improve. This would be helpful for treatment planning and policy decisions regarding how much therapy is enough. The progress of 125 clients entering psychological treatment in two university training clinics was tracked on a session-by-session basis using a naturalistic design. Recovery and reliable improvement were the bi-nomial events of interest used in a survival analysis that estimated the number of weekly treatment sessions needed to meet criteria. Using survival analysis it was estimated that it would take about 8 sessions for 50% of clients to show reliable improvement and 21 sessions for about 85% to meet this criterion. Recovery took more treatment, with 50% of clients estimated to recover after 14 sessions and 70% requiring 23. On the basis of the present results we conclude that the present policy of the Australian Government in both the public and private sector regarding the number of sessions needed for clients entering psychological treatments to show a benefit is much less than is, in fact, necessary. The findings of the current study are roughly consistent with those found elsewhere and suggest a minimum benefit should be closer to 20 sessions. The current policy appears to be suitable for only about one-third of clients who carry the burden of psychological illness.

Keywords: Psychotherapy outcome, dose response, effects of psychotherapy, psychological treatment, survival analysis

Introduction

Across the developed world government bodies, agencies, and insurance companies frequently limit the number of sessions available to clients undergoing psychological treatments. The reasons for doing so are often for cost saving purposes and sometimes to ration care, making it available for more individuals and thereby shortening the time clients are kept on waitlists. Restricting services by limiting financial support on the basis of sessions, although common, is highly variable from place to place and seldom based on empirical data on the need for treatment. Rather, specific limits are set as if they were appropriate and reasonable for most individuals. In the United States for example it is common practice for managed healthcare companies to limit sessions to 4, 8, or with more generous plans, perhaps as many as 20 sessions per year. Data from the directors of 73 United States' college and university counselling centres shows that more than half of these centres have a self-imposed session limit of 10 sessions (Stone & McMichael, 1996). In Germany private insurers often limit sessions to 40 per year for practitioners of cognitive behaviour therapy and 60 or 80 per year for those seeing a practitioner trained in psychodynamic therapy, even though there do not appear to be differences in the degree of disturbance or outcome for clients entering these treatments. With the recent Medicare initiative in Australia, clients can get 6 sessions with a psychologist or clinical psychologist and then another 6 after review within a calendar year if they are referred by a general practitioner (GP). To be referred by a GP they need a Mental Health Care Plan containing a diagnosis, an activity that would encourage the therapist to use at least one of the treatment sessions for the purpose of assessment and diagnosis.

Session limit policies such as these seem to arise without public discussion of the consequences or empirical data to justify such actions. Available data on how much psychotherapy is enough is somewhat limited but consistently suggests more is better than less (at least with regards to group averages) (Consumer Reports, 1995, November; Johnson, 1965; Seeman, 1954; Standal & van der Veen, 1957; Strassberg, Anchor, Cunningham, & Elkins, 1977; Weitz, et al., 1975). If on average, more therapy is better, how much therapy is enough? In recent years psychotherapy outcome research has addressed this question using survival analysis or related techniques to model the point in therapy a clinically significant change occurs for each individual client (Anderson & Lambert, 2001; Hansen, Lambert, & Forman, 2002; Kadera, Lambert, & Andrews, 1996; Kopta, Howard, Lowry, & Beutler, 1994). Such research requires measurement of client dysfunction on a session-by-session

basis and a binomial classification of dysfunction (recovered, not recovered) at each session. The results from these studies suggest a range of 11–18 sessions as the number of sessions needed to achieve clinically significant change for half of the sample starting therapy. Typically around 20% of these clients also receive psychoactive medications.

These four studies used samples from various sites including an outclient mental health centre staffed by graduate student psychotherapists (Anderson & Lambert, 2001; Kadera et al., 1996), a sample of outclients seen at five mental health centers (Kopta et al., 1994), and a large multi-site sample comprised mostly (70%) of clients receiving Employee Assistance Services with the rest coming from private practice in America (Hansen & Lambert, 2003). It is not clear if these estimates generalise to treatment settings in other countries. Such data could be especially beneficial to clinicians who treat such groups and to those who make decisions about allocating treatment or referral for such treatment.

The present study was undertaken in order to provide information on the dose-response relationship in psychotherapy with the hope that it will shed some light on the consequences of dosage restrictions for client well being.

Method

Participants

The client sample was composed of 125 adults (18

years old and older) who presented for individual psychotherapy at either the Griffith University or University of Queensland Psychology Clinics. These clinics offer reduced fee services to the surrounding community and are accredited training facilities in clinical psychology. Clients are eligible to be seen for as few sessions or as many sessions as deemed necessary by their therapist and the therapist's supervisor. Clients presented with a wide range of problems, representative of the population that utilises clinical psychology services. Most were primarily experiencing a mood, anxiety, or adjustment disorder. The rest of the diagnoses were inclusive of many other DSM-IV diagnoses (e.g., personality disorder, eating disorder, substance abuse, interpersonal conflict, bereavement, sleep disorder, sexual dysfunctional).

The client pool for this study included individuals who were seen at least twice for therapy for whom session data was available. A total of 132 clients were identified who presented at least twice for treatment. Of this initial pool seven clients were excluded from the final data analysis because baseline data was not available. Demographic information was collected from all clients upon intake. The age of the clients ranged from 18–65, with a mean age of 34.5 years (SD¹/₄12.3). Sixty-five percent of the sample was female and 35% were male. The mean baseline score on the measure of psychological functioning (OQ- 45) was 79 (SD¹/₄25), with a range from 7–126. This score is typical for persons receiving outclient psychotherapy. Ninety-one (72.8%) clients reported a score above the clinical cutoff at baseline.

A total of 77 trainee psychotherapists enrolled in post-graduate training in clinical psychology participated in this research. All psychotherapists were registered or provisionally registered psychologists with the state registration board. Eighty-two percent were female and 18% were male. The majority of the therapists were Caucasian. The student psychotherapists endorsed a variety of therapeutic orientations, with cognitive-behavioral, interpersonal, and psychodynamic as the major theories from which they practiced most frequently.

The outcome measure used was the Outcome Questionnaire-45 (OQ-45) (Lambert et al., 2004). The OQ-45 is a 45- item, self-report measure of psychological functioning that calls on clients to reflect on how they have been feeling and functioning during the prior week. Items are measured on a 5- point Likert scale: 0¹/₄never, 1¹/₄rarely, 2¹/₄sometimes, 3¹/₄frequently, 4¹/₄almost always. The range of scores possible on the OQ-45 is 0– 180, with higher scores reflecting more severe distress. The OQ-45 assesses three broad domains of client functioning: (1) subjective discomfort (e.g., mainly symptoms of anxiety and depression—''I feel blue''), (2) interpersonal relationships (e.g., ''I feel lonely''), and (3) social role performance (e.g., ''I have too many disagreements at work/school''). The development of norms for the OQ-45 was based on data collected in the USA (Lambert et

al., 2004; Umphress, Lambert, Smart, Barlow, & Clouse, 1997). Previous studies have provided information about the psychometric properties of the OQ-45.

The internal consistency of the OQ-45 approximates .93, and 3-week test-retest reliability hovers around .84 (Lambert et al., 2004). Concurrent validity of the OQ-45 has been demonstrated through correlates with a variety of scales and across several countries (e.g., Symptom Checklist-90-R, r¹/₄.78 (Derogatis, 1983); Beck Depression Inventory, r¹/₄.80 (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). It appears to be a solid measure of psychopathology. To measure and validate the meaning of therapeutic change and to cast it as a binomial event, clinically significant change (CS) has become widely used in psychotherapy research (Jacobson, Roberts, Berns, & McGlinchey, 1999; Kendall, 1999). The CS criteria as proposed by Jacobson and colleagues require two components in order for change during psychotherapy to be considered meaningful (Derogatis, 1983; Jacobson et al., 1999). The first criterion maintains that a client must improve beyond a cutoff score set to distinguish a functional sample from a dysfunctional sample. When a client's OQ-45 score moves from the dysfunctional range of scores past the cut-off and therefore into the range of functional scores the client is considered to be functioning in the normal range, i.e., to be considered meaningful, the functioning of the client when therapy is complete is more similar to the functional group than a sample of clients with known psychopathology.

The second criterion maintains that the change across time must be a statistically reliable change (RC) (i.e., the change must be greater than the measurement error of the outcome instrument employed). In the present study, Jacobson and colleagues' (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991) CS criteria for the OQ-45 were used to evaluate the final outcome for each client and the point at which this phenomenon occurred. The OQ-45 Administration and Scoring Manual (Lambert et al., 2004) provides information on both criteria. To satisfy both criteria, a client's change was regarded as clinically significant when the final OQ-45 score was equal to or less than 63, and had decreased by at least 14 points from the intake session. For those clients who began treatment at or below an OQ-45 score of 63 only reliable change (RC) could serve as the criterion for improvement. This single criterion (14b points change) could also be applied to those clients who began treatment in the dysfunctional range.

Support for the validity of the OQ-45's reliable change and clinical significance cutoff scores have been reported and replicated by (Bauer, Lambert, & Nielsen, 2004; Beckstead, et al., 2003; Lunnen & Ogles, 1998). The Jacobson/Truax formulas provide a reasonable basis for estimating cutoff scores and classification of change that show considerable consensus with other self report measures such as the Symptom Checklist List-90 and Beck Depression Inventory.

In 2006 and 2007, clients attending an intake or an individual psychotherapy session were given an OQ- 45 by the receptionist when they checked in for their appointment prior to each session. Clients also completed an OQ- 45 prior to each subsequent session of therapy. For the purposes of this study, the OQ45 scores were not provided to therapists during the course of therapy. The OQ-45 total score recorded before or during the intake session was considered the baseline for that client. In order to estimate dose-response the OQ-45 total score recorded before the second session represents the effects of one session of therapy, the OQ-45 score recorded before the third session represents the effects of two sessions and so on.

Survival analysis

Clients attended between 2 and 34 sessions. The mean number of sessions attended was 9.5 (median ¹/₄8, mode¹/₄3). Sixty-four percent of the sample attended 10 sessions or fewer. Five percent of the sample attended 20 sessions or more. The mean number of days in therapy (calculated as the number of days elapsed between the baseline and final sessions) was 95.1 days (13.6 weeks) with a median of 81 days (11.6 weeks) and mode of 28 days (4 weeks). A total of 1,034 OQ-45s were collected (88.7% of the total number of sessions). While OQ- 45s were not collected on 11.3% of the sessions, the statistical procedures employed use statistical modelling to deal with this missing data.

The Kaplan-Meier (Kaplan & Meier, 1958) procedure for survival analysis was applied to the data from this study. The Kaplan-Meier procedure has been called the foremost nonparametric technique for analysing survival data (Collett, 1994; Cox & Oakes, 1984; Greenhouse, Stangle, & Bromberg, 1989; Willet & Singer, 1993). This procedure estimates the proportion of the population who will attain CS (or RC) at session one, session two, and so on. The Kaplan-Meier procedure as used in this study does this by analysing the raw data of sessiontosession change with CS (or RC) criteria as the dependent variable. When a client improves by 14 points without ever backsliding at a latter session that session is considered the point at which reliable change (RC) occurred. When the client improves by 14 points and crosses the cut-off of 64/63, CS is regarded to have occurred at that session (provided no back sliding occurred). Thus, a client who ends therapy and who is not changed according to RC/CS criteria at termination is not counted as meeting criteria even though at some point in treatment they may have. As some clients never achieve RC or CS change before ending therapy their data is said to be "censored". Censored data are not discarded from the survival analysis but used in modelling the treatment dosage model. Three survival analyses were completed using only the RC criterion, one for the "functional" sample (initial OO-45 score below 64), one for the "dysfunctional" sample (initial OO-45 score 64 or above), and one for the full sample. In addition, a survival analysis was performed employing the two criteria, i.e., CS. Since only those clients who began treatment in the dysfunctional range could pass the threshold for normal functioning, this was carried out on the "dysfunctional sample only".

Results

The first survival analysis was carried out for RC using the number of sessions as the time variable. The analysis was carried out with the whole sample (i.e., the "functional" and "dysfunctional" samples combined). Of the 125 clients, 64 (51.2%) achieved RC at some stage of therapy. This analysis estimated that 5.6% of the clients achieved RC by session 2, 20.4% of the clients achieved RC by session 4, 50.6% of the clients achieved RC by session 10, and 76.1% of the clients had achieved RC by session 18. Figure 1 displays a graphic representation of cumulative RC probability data from this analysis. This survival analysis was repeated separately for the 91 "dysfunctional" and 34 "functional" clients. Looking first at the "dysfunctional" clients, the analysis estimated that 15.0% of the clients achieved RC by session 3, 35.1% of clients achieved RC by session 6, 48.9% achieved RC by session 8, and 83.7% achieved RC after 21 sessions. Looking next at the "functional" clients, the analysis estimated that 15.2% of the clients achieved RC by session 5, 34.4% of clients achieved RC by session 9, 43.5% achieved RC probability data from this analysis.

The next survival analysis was carried out for CS using the 91 "dysfunctional" cases. Of the 91 "dysfunctional" clients, 31 (34.1%) were estimated to have achieved CS at some stage of therapy. This analysis found that 8.8% of clients were estimated to achieve CS by session 4, 34.7% by session 10, 50% CS by session 14, and approximately 70% were estimated to have achieved CS by session 23. Figure 3 displays a graphic representation of cumulative CS probability data from this analysis.



Figure 1. Kaplan-Meier survival plot for clients of whole sample who reach reliable change.







Figure 3. Kaplan-Meier survival plot of dysfunctional clients who reach clinical significance.

Discussion

Even a casual observer of the diversity of policy and practices concerning treatment length, worldwide, might be quite puzzled at the absence of the use of empirical data to shape important policy decisions concerning treatment length. A central goal of this study was to examine the relationship between treatment sessions and outcome in order to contribute empirical knowledge for social policy regarding setting session limits for clients undergoing psychological treatments. The results suggested that *reliable change* was estimated to be achieved for 50% of those who were treated, if clients were to undergo 10 sessions of psychotherapy. For 50% of clients to achieve *clinically significant change* an estimated 14 sessions were required. About 70% of clients who enter treatment with moderate or more levels of disturbance would achieve clinically significant change if they remained in treatment for 23 sessions. These estimates are fairly consistent with estimates coming out of the USA; the range for 50% of clients to attain clinically significant change is about 11–18 sessions (Anderson & Lambert, 2001; Hansen et al., 2002; Kadera et al., 1996; Kopta et al., 1994).

It should be recalled that in all these studies clinically significant change is modelled and that the actual outcome of clients is less than the predicted outcome because a subset of clients leave treatment before this level is reached. In the present study about 34% of clients actually met criteria for clinically significant change; a proportion that is also consistent with past research. These data suggest the need for providers and service managers to give high priority to keeping clients in treatment longer, rather than setting session limits or reimbursement practices for shorter durations, if the goal of an intervention is to return clients to within a standard deviation of normal functioning.

Obviously, the recent Medicare initiative in Australia where GP-referred clients can get six sessions and then another six (within a calendar year after review and approval) falls far short of the amount of treatment necessary for most clients to receive a substantial benefit. It should be noted here that many organisations claim to be delivering "evidence-based" services while ignoring the fact that most outcome studies that supply the evidence (clinical trials) are based on treatment lengths that hover around 14 weeks (Hansen & Lambert, 2003). Often such studies, especially if they offer CBT, include closer to 20 sessions because they require two sessions weekly in the first weeks of therapy. It is obvious that 6–12 sessions of psychotherapy is not consistent with evidence-based practice recommendations, nor the practice-based evidence provided in this study. Such guidelines leave at least half the clients whose services are reimbursed 50% underserved. This is an especially problematic when one considers the burden of illness carried by those who have mental health problems and the financial costs (in the form of lost productivity and absenteeism) of such illnesses to society.

An important issue to be considered in generalising the results of this study is the degree to which dose response curves generated on the basis of trainee-delivered psychotherapy are representative of those that would be generated based on routine care. The psychotherapy outcome literature has been consistent in showing little difference in outcomes between professional and degreed providers (Beutler, Machado, & Neufeld, 1994; Stein & Lambert, 1995), with some studies showing superior outcomes for trainees (M. Lambert et al., 2003). A number of factors have been thought to account for these findings, including the fact that trainees have much smaller case loads with their ongoing activities being monitored and supervised by professionals. Hansen and Lambert (2003) found no difference in the dose response curves of professionals and students in training.

Although setting general session limits has great advantages for managing costs and making the life of administrators easier, it is not to the advantage of most clients and diminishes the overall value of services, especially for the most disturbed clients. We interpret the dose-response data reported in the present study as indicating that if a general session limit must be set it would be more reasonable to think in terms of 20–25 sessions for most clients, perhaps depending on initial severity of disturbance. The highest quality of care (and wisest financial decision) will be provided by monitoring each client's treatment response and making treatment length (number of sessions) a function of treatment response and mental health status, rather than arbitrary or theory-driven limits. Termination of services would be instigated only when it becomes clear that no further progress can be expected if a client continues to receive the same treatment. This course of action is probably justified because the survival curves generated in the current study (like those from most past studies) show a

tendency to flatten as the number of sessions goes up, with few clients meeting criteria for reliable change and clinically significant change at subsequent sessions following the 25th session, six months of weekly treatment. It should be kept in mind that this study reported on data where feedback on progress and deviation from normal functioning was withheld from therapists and clients. Feedback to therapists on client progress can result in reductions in the cost of treatment. Where client progress is routinely monitored, clients who respond rapidly to treatment will depart from therapy with fewer sessions (M. J. Lambert et al., 2003) providing savings that will offset the costs of providing more sessions to patients who are responding more slowly to treatment (Barkham, Hardy, & Mellor-Clark, 2010). Thus, our recommendation that client's will be better served by providing a dose of therapy based on routine monitoring of the clients' response to therapy will not necessarily be more expensive.

References

Anderson, E., & Lambert, M. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology*, 57, 875–888.

Barkham, M., Hardy, G. E., & Mellor-Clark, J. (2010). Developing and delivering practice-based evidence. Chichester, West Sussex: John Wiley & Sons.

Bauer, S., Lambert, M., & Nielsen, S. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality* Assessment, 82, 60–70.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 561–571.

Beckstead, D., Hatch, A., Lambert, M., Eggett, D., Goates, M., & Vermeersch, D. (2003). Clinical significance of the Outcome Questionnaire (OQ-45.2). *The Behavior Analyst Today*, 4, 79–90.

Beutler, L., Machado, P., & Neufeld, S. (1994). Therapist variables. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.

Collett, D. (1994). Modeling survival data in medical research. New York: Chapman and Hall.

Consumer Reports (1995, November). Mental Health: Does therapy help? (November), 734-739.

Cox, D., & Oakes, D. (1984). Analysis of survival data. New York: Chapman and Hall.

Derogatis, L. R. (1983). The SCL-90: Administration, scoring, and procedures for the SCL 90. Baltimore: Clinical Psychometric Research.

Greenhouse, J., Stangle, D., & Bromberg, J. (1989). An introduction to survival analysis. *Journal of Consulting and Clinical Psychology*, 57, 536–544.

Hansen, N., & Lambert, M. (2003). An evaluation of the doseresponse relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research*, 5, 1–12.

Hansen, N., Lambert, M., & Forman, E. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9, 329–343.

Jacobson, N., Follette, W., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinically significance. *Behavior Therapy*, 15, 336–352.

Jacobson, N., Roberts, L., Berns, S., & McGlinchey, J. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology*, 67, 300–307.

Jacobson, N., & Truax, P. (1991). Clinically significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.

Johnson, R. (1965). Number of interviews, diagnosis, and success of counseling. Journal of Counseling Psychology, 12, 248–251.

Kadera, S., Lambert, M., & Andrews, A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy: Practice and Research*, 5, 1–20.

Kaplan, E., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53, 457–481.

Kendall, P. C. (1999). Clinical Significance. Journal of Consulting and Clinical Psychology, 67, 283-285.

Kopta, S., Howard, K., Lowry, J., & Beutler, L. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 1009–1016.

Lambert, M., Morton, J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R., et al. (2004). Administration and scoring manual for the Outcome Questionnaire-45. SLC, UT: OQ Measures.

Lambert, M., Whipple, J., Hawkins, E., Vermeersch, D., Nielsen, S., & Smart, D. (2003). Is it time for clinicians to routinely track patient outcome?: A meta-analysis. *Clinical Psychology: Science & Practice*, 10, 288–301.

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

Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology*, 66, 400–410.

Seeman, J. (1954). Psychotherapist judgments of therapeutic process and outcome. In C. R. Rogers & R. F. Dymond (Eds.), *Psychotherapy* and personality change (pp. 99–108). Chicago, IL: University of Chicago Press.

Standal, S., & van der Veen, F. (1957). Length of therapy in relation to psychotherapist estimates of personal integration and other case variables. *Journal of Consulting Psychology*, 21, 1–9.

Stein, D., & Lambert, M. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced? *Journal of Consulting and Clinical Psychology*, 63, 182–196.

Stone, G., & McMichael, J. (1996). Thinking about mental health policy in university and college counseling centers. *Journal of College Student Psychotherapy*, 10, 3–27.

Strassberg, D., Anchor, K., Cunningham, J., & Elkins, D. (1977). Successful outcome and number of sessions: When do psychotherapists think enough is enough? *Journal of Counseling Psychology*, 24, 477–480.

Umphress, V., Lambert, M., Smart, D., Barlow, S., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. Journal of Psychoeducational Assessment, 15, 40–55.

Weitz, L., Abramowitz, S., Steger, J., Calabria, F., Conable, M., & Yarus, G. (1975). Number of sessions and client-judged outcome: The more the better? *Psychotherapy: Theory, Research and Practice*, 12, 337–340.

Willet, J., & Singer, J. (1993). Investigating onset, cessation, relapse, and recovery: Why you should, and how you can use discrete-time survival analysis to examine event occurrence. *Journal of Consulting and Clinical Psychology*, 61, 952–965.