Cost-Effectiveness and Cost-Utility of Cognitive Therapy, Rational Emotive Behavioral Therapy, and Fluoxetine (Prozac) in Treating Depression: A Randomized Clinical Trial

Florin A. Sava
Babes-Bolyai University, Cluj-Napoca, Romania

Brian T. Yates
American University

Viorel Lupu
Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania

Aurora Szentagotai and Daniel David
Babes-Bolyai University, Cluj-Napoca, Romania

Cost-effectiveness and cost-utility of cognitive therapy (CT), rational emotive behavioral therapy (REBT), and fluoxetine (Prozac) for major depressive disorder (MDD) were compared in a randomized clinical trial with a Romanian sample of 170 clients. Each intervention was offered for 14 weeks, plus three booster sessions. Beck Depression Inventory (BDI) scores were obtained prior to intervention, 7 and 14 weeks following the start of intervention, and 6 months following completion of intervention. CT, REBT, and fluoxetine did not differ significantly in changes in the BDI, depression-free days (DFDs), or

This research was supported by the Albert Ellis Institute, the National Council for Research (Romania), and the Romanian Center for Cognitive and Behavioral Psychotherapies (Grants 113, 33374, and 9, respectively). These organizations had no role in the design and implementation of the study. Details of procedures and efficacy findings are provided in David (2006).

Correspondence concerning this article should be addressed to: Brian T. Yates, Department of Psychology, American University, 4400 Massachusetts Avenue, N.W., Washington, DC 20016–8062; e-mail: brian.yates@mac.com
Quality-Adjusted Life Years (QALYs). Average BDI scores decreased from 31.1 before treatment to 9.7 six months following completion of treatment. Due to lower costs, both psychotherapies were more cost-effective, and had better cost-utility, than pharmacotherapy: median $26.44/DFD gained/month for CT and $23.77/DFD gained/month for REBT versus $34.93/DFD gained/month for pharmacotherapy, median $/QALYs = $1,638, $1,734, and $2,287 for CT, REBT, and fluoxetine (Prozac), respectively. © 2008 Wiley Periodicals, Inc. J Clin Psychol 65:36–52, 2009.

Keywords: depression; cost-effectiveness; cost-utility; pharmacotherapy; CBT; REBT; fluoxetine; DFD; QALY; psychotherapy

Antidepressant medication is promoted as the “gold standard” for treatment of major depressive disorder (MDD; American Psychiatric Association, 2000; Olfson & Klerman, 1993) despite experiments and reviews suggesting that evidence-based psychotherapies are at least as effective as pharmacotherapy in treating MDD (e.g., Antonuccio, Danton, & DeNelsky, 1995; Revicki, Siddique, Frank, Chung, Green, Krupnick et al., 2005; Schulberg, Raue, & Rollman, 2002; Scott, Palmer, Paykel, Teasdale, & Hayhurst, 2003). This may be due to an assumption that medication costs less than psychotherapy, which need not be the case. Costs of psychotherapies are measured correctly according to the duration of therapy and follow-up visits, but costs of pharmacotherapy are underestimated if limited to doses and monitoring services delivered over the same period for which psychotherapy was provided (e.g., 12 weeks), given that pharmacotherapy may continue for many months more, possibly through the lifetime of the individual. In addition, if only total monetary costs are reported, differences found in costs may be due to variation between countries in the availability—and thus costs—of therapists relative to therapeutic drugs.

Psychotherapy researchers studying cognitive-behavioral therapies (CBTs) seem reluctant to measure, analyze, and report costs, benefits, and cost-effectiveness (Hunsley, 2003; Prigatano & Pliskin, 2003; Yates, 1994). This would be inconsistent with high-profile endorsements to consider costs, cost-effectiveness, and cost-benefit among the evidence supporting use of a particular therapy. The American Psychological Association (APA)’s Presidential Taskforce on Evidence-Based Practice (2006), for example, reported that “APA endorses multiple types of research evidence (e.g., efficacy, effectiveness, cost-effectiveness, cost-benefit, epidemiological, treatment utilization) that contribute to effective psychological practice.” (p. 274). It is, in fact, APA policy since 2002 that evidence on clinical utility:

... at a minimum ... includes attention to generality of effects across varying and diverse patients, therapists, settings, and the interaction of these factors; the robustness of treatments across various modes of delivery; the feasibility with which treatments can be delivered to patients in real-world settings; and the costs associated with treatments [italics added]. (p. 275).
Measurement and reporting of costs and benefits—critical determinants of reimbursement policy—may have been left to researchers who consider psychotherapy as at best an adjunctive, and not an alternative, treatment for depression. Inconsistent measurement of costs, and of outcomes, also may prevent conclusive statements about whether psychotherapy or pharmacotherapy is more effective, less costly, or both, relative to each other and relative to usual care.

Additionally, although most studies use a common measure of outcome, the Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979), that measure may not be as relevant to policy makers and funders as Depression-Free Days (DFDs; Lave, Frank, Schulberg, & Kamlet, 1998) and Quality-Adjusted Life Years (QALYs; Gold, Siegel, Russell, & Weinstein, 1996; Weinstein & Stason, 1977). Worse yet, the quality of research design also has ranged from randomized clinical trials to self-assignment to psychotherapy or pharmacotherapy (Gabbard, Lazar, Hornberger, & Spiegel, 1997).

Those problems in measurement and design may explain why, relative to studies that compare costs of different pharmacotherapies for depression, reviews that compare costs of pharmacotherapies versus psychotherapies seem more likely to report conflicting findings. For example Schulberg et al. (2002) and Pirraglia, Rosen, Hermann, Olchanski, and Neumann (2004) concluded that CBT and interpersonal psychotherapy both were more costly and more effective than were a variety of control and “usual care” interventions, after reviewing nine and 13 studies, respectively. Pirraglia et al. also found that pharmacotherapies for depression had lower costs per QALY than did psychotherapies for depression, and therefore better cost-utility (Levin & McEwan, 2001). However, a review by Barrett, Byford, and Knapp (2005) of 16 studies concluded that pharmacotherapies and psychotherapies have similar cost-effectiveness in treating depression.

Antonuccio, Thomas, and Danton (1997) found that the effectiveness of psychotherapies and pharmacotherapies was sufficiently similar to focus the examination of cost-effectiveness exclusively on possible differences in cost. Antonuccio et al. (1997) estimated that over a 2-year period, pharmacotherapy (fluoxetine or Prozac) resulted in 23% higher costs than individual CBT while also estimating that group CBT produced only a 2% cost savings relative to individual CBT. Using a simulation modeling technique, Haby, Tonge, Littlefield, Carter, and Vos (2004) concluded that CBT provided by psychologists working in public health settings was more cost-effective than was pharmacotherapy using serotonin reuptake inhibitors (SSRIs) for MDD in children and adolescents. Haby et al. also found that CBT provided by psychologists working in private settings was similar in cost-effectiveness to pharmacotherapy using selective SSRIs. Cost-effectiveness data from more empirical studies involving CBT were inconclusive. Revicki et al. (2005) found no difference in cost-effectiveness between CBT and pharmacotherapy with paroxetine (Paxil) in a randomized clinical trial using a U.S. sample of 267 low-income, minority women. Bower et al. (2000), too, found no difference in cost-effectiveness between CBT, nondirective counseling, and usual care provided by a general practitioner, with a sample of 464 patients presenting depression or mixed anxiety and depression.

In an effort to resolve these findings regarding the effectiveness, costs, and cost-effectiveness of pharmacologic and psychotherapeutic interventions for depression, we assessed and analyzed both the costs and effectiveness of a pharmacotherapy and of two cognitive therapies. Clients diagnosed with MDD were assigned randomly to either (a) SSRI pharmacotherapy (fluoxetine, Prozac) or (b) the most common
therapy for depression, Cognitive Therapy (CT; see Beck, 1976) or (c) another increasingly common therapy, Rational Emotive Behavior Therapy (REBT; Ellis, 1987). As with many other clinical trials for depression treatments (e.g., the National Institute of Mental Health Treatment of Depression Collaborative Research Program; Elkin et al., 1989), and following the American Psychiatric Association (2000) guidelines for treatment of MDD, we excluded a placebo treatment condition. Butler, Chapman, Forman, and Beck’s (2006) review of methodologically rigorous meta-analyses of CBT showed that it has substantial effect sizes well above those found for no-treatment and waiting-list conditions. Butler et al. also noted that “CBT was somewhat superior to antidepressants in the treatment of adult depression” (p. 17).

Choice of Treatments for Depression

We chose fluoxetine (Prozac) for the pharmacotherapy because it is one of the most commonly prescribed antidepressants and is often described as “cost-effective.” Whereas Chisholm, Sanderson, Ayuso-Mateos, and Saxena (2004) found that older tricyclic antidepressants (TCAs) such as imipramine could be more cost-effective than SSRIs—especially in lower income regions—Barrett et al. (2005) concluded that fluoxetine (Prozac) and other SSRIs appear consistently more cost-effective than most TCAs for a variety of patients. Fluoxetine (Prozac) also appears to be safer than TCAs, in that fluoxetine produces fewer side effects (Mulrow et al., 2000, as cited in Serrano-Blanco et al., 2006) than TCAs and has become relatively inexpensive (Serrano-Blanco et al., 2006).

We chose Beck’s CT, a common form of CBT, because it is a widely used psychotherapy for depression with a considerable evidence base (e.g., American Psychiatric Association, 2000; DeRubeis, Gelfand, Tang, & Simons, 1999; Dobson, 1989; Hollon, Shelton, & Davis, 1993). CT includes behavioral activation and dysfunctional thought modification as well as identification and structural modification of generalized core beliefs presumed to be principle causes of depressive reactions and dysfunctional thinking. We chose a second form of CBT, REBT, because it focuses on different cognitive processes (Ellis, 1987), and it also is accumulating evidence of its effectiveness (Engels, Garnefsky, & Diekstra, 1993). REBT attempts to change cognitive processes by (a) advocating unconditional self-acceptance; (b) focusing explicitly on reducing secondary problems such as depression about depression; (c) focusing on demandingness, which seems to be the core belief involved in MDD; and (d) emphasizing the proximal causes of negative feelings (i.e., irrational beliefs).

Comparisons of Effectiveness and Costs

Our study used a randomized clinical-trial design, measured adherence to treatment protocols, and examined and reported attrition rates for each treatment. A separate article (David, Szentagotai, Lupu, & Cosman, 2008) has detailed the design, protocols, and some effectiveness findings, written according to Consolidated Standards of Reporting Trials, or CONSORT, criteria (Moher, Schulz, & Altman, 2001). To facilitate comparison of our findings with other research on psychotherapies and pharmacotherapies for depression and for other mental illnesses, we report BDI, DFD, QALY, cost per QALY gained ($/QALY), and cost per DFD gained per month ($/DFD).
The validity of most cost findings for depression treatments may be limited by confinement of almost all of this research to four of the most developed countries on Earth: Australia, Canada, the United Kingdom, and the United States of America. Adjustments for differences between nations in currency and cost of living, as well as inflation or deflation since the year cost data were collected, are relatively minor issues. Of more concern are variations between countries in the availability and price of therapist time. Costs of therapist time may be lower in other countries, even after monetary adjustments, because providing therapy may not be valued as highly as are other professions.

To our knowledge, only one study on the costs as well as the effectiveness of alternative treatments for depression has been conducted in a country not as developed as the aforementioned “big four.” In Chile, psychoeducational treatment by social workers was found to be more cost-effective than was pharmacotherapy for depression (Araya, Flynn, Rojas, Fritsch, & Simon, 2006). We attempted to test the external validity of both this and similar findings reviewed earlier by conducting our research in a country that is 1 of 40 “upper middle income economies,” according to the World Bank (2005). Included in these countries are Argentina, Mexico, Poland, Russia, and the nation in which this study was conducted: Romania. Inducted into the European Union in January 2007, Romania has a per capita Gross Domestic Product of $9,446. The average gross wage per month in Romania is $545.36 (1418 RON or €419.38) based on international exchange rates when data were collected, and $846.06 based on purchasing power parity when data were collected (International Monetary Fund, 2006).

We attempted to improve the potential replicability of cost and cost-effectiveness findings reported here by itemizing costs, so that the type and amount of each “ingredient” resource are available in sufficient detail for other researchers to examine and convert into current and locally meaningful numbers (Yates, 1996). Costs were itemized for providers, clients, facilities, and materials. The monetary value of each specific resource was obtained for the local economy, as were the amounts and types of each resource used by each treatment. To make our cost findings potentially usable by more interest groups involved in funding decisions, we itemized time spent by consumers as well as providers in treatment and treatment-related activities. Readers who wish to exclude these costs from cost analyses can readily do so, given the breakdowns provided for time and costs for providers, consumers, and a variety of other treatment resources.

Method

Participants

Calculations for statistical power indicated that $n$s of 52 would detect the medium effect size expected from Jacobson et al. (1996) and our experience in clinical trials, with power at .80, $\alpha = .05$, as detailed in David (2006; the manual of the trial published in Romania). The 170 adult participants (115 female, 55 male) with no comorbidity met criteria for MDD according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; American Psychiatric Association, 1994), scored at least 20 on the BDI, and scored at least 14 or higher on the Hamilton Rating Scale for Depression (Hamilton, 1967). The support for such a standard procedure of selection is presented elsewhere (DeRubeis et al., 1999; Hollon et al., 1993). As detailed in David (2006) and David et al. (2008), attrition during treatment was 14% in the medication condition, 10% in CT, and 9% in
REBT. Somewhat higher attrition in the medication condition seemed due to side effects of medication (i.e., nausea, sexual dysfunction). These dropout rates may be lower than usual due to the use of a different population in a different healthcare system than is usually the case in published studies, but other researchers have reported reasonably comparable attrition rates of 15% for CBT and 16% for medication (DeRubeis et al., 2005).

Analyses of variance found no significant differences in the following demographic variables between the three treatment conditions. Participants were a mean 35, 39, and 37 years old in the REBT, CT, and medication conditions (range = 21–57), respectively. Approximately half had university educations (30, 28, and 31 of the 57, 56, and 57 participants in the REBT, CT, and medication conditions, respectively); the remainder had high-school educations. Participants were recruited from four public clinics and research institutes, private practices, and by public-service announcements. Participants were not paid but received the therapy to which they had been assigned free of charge. This study was funded by a combination of three private and public research organizations, based in both Romania and the United States, as detailed in the footnotes. Additional demographics and recruitment information beyond the scope of this article are provided by David et al. (2008).

For persons eligible for participation, the number of previous episodes of depression was 3.6, 4.0, and 3.9 for the REBT, CT, and medication conditions, respectively. Persons with concurrent psychiatric disorders (i.e., bipolar or psychotic subtypes of depression, panic disorder, current substance abuse, past or present schizophrenia or schizophreniform disorder, organic brain syndrome, and mental retardation; also see Jacobson et al., 1996) were excluded from the study, as were those already receiving psychotherapy, psychotropic medication, or hospitalization due to imminent suicide potential or psychosis.

Participants qualifying for the study provided written informed consent prior to random assignment to one of the three conditions [i.e., CT, REBT, or fluoxetine (Prozac)]. The pharmacotherapy condition served as the reference condition in this study because it is the “golden standard treatment” for depression according to the American Psychiatric Association (2000). Statistical tests found no significant differences prior to treatment in the three conditions, in terms of number of previous episodes of depression, presence or absence of dysthymia, severity of depression, gender, or marital status (for details, see David, 2006).

Procedure

A large number of patients were evaluated for the study from 2001 to 2004. To assess their eligibility, diagnoses were determined by the Structured Clinical Interview for DSM-IV (First, Spitzer, Gibbon, & Williams, 2001). A clinical psychologist then evaluated potential participants, the first BDI was administered, and participants were assigned to conditions. At the middle and the end of the intervention (Weeks 7 and 14) and at a 6-month follow-up, participants were administered the BDI by a research assistant unaware of the condition to which the participant had been assigned.

Therapists. Psychotherapies were provided by eight experienced therapists (six psychologists, two psychiatrists), all certified in CT and REBT, averaging 11 years (range = 7–14) of postdegree clinical experience. These therapists’ degrees included the Romanian equivalents of master’s (n = 4) and doctoral (n = 2) degrees for the psychologists, and the doctor of medicine degree for the two psychiatrists. All had
practiced for at least 7 years following formal training. Pharmacotherapy was provided by four psychiatrists with an average 16 years (range = 14–18) of postdegree clinical experience. Manuals elaborated guidelines for each of the three interventions. Sessions were audiotaped at random. Tapes were checked for protocol violation as well as the quality of intervention, finding high levels of fidelity in all three conditions (David, 2006).

CT. CT was based on the techniques and descriptions in the Beck et al. (1979) and Beck (1976) manuals. Sessions focused on the identification and structural modification of automatic thoughts and generalized core beliefs (schemas) that are presumed to be the major causes of dysfunctional thinking and depressive reactions (Dowd, 2006; Szentagotai et al., 2006).

REBT. REBT was based on manuals developed by Ellis and others (Ellis, 1962/1994; Walen, DiGiuseppe, & Dryden, 1992). Treatment focused on the irrational beliefs hypothesized to mediate depressive symptoms: demandingness and self-downing. Other core irrational beliefs were approached if identified.

CT, REBT, and pharmacotherapy involved a maximum of 20 individual 50-min therapy sessions or a maximum of 20 individual 20-min pharmacotherapy sessions, over the 14 weeks allowed for treatment. Clients received an average of 17.5 sessions of CT, 18.0 sessions of REBT, or 18.0 sessions of fluoxetine monitoring.

Fluoxetine (Prozac). Over 14 weeks of treatment, participants assigned to the fluoxetine (Prozac) condition attended one session per week with a psychiatrist. To better approximate real-world practice, these sessions were kept shorter than the CBT sessions: After an initial 50-min meeting with the psychiatrist, subsequent sessions were 20 min. Although this may be a longer session duration than is common practice in the United States, it is not radically different from the 20- to 30-min biweekly session reported in the literature for clinical research on fluoxetine in some populations (e.g., Treatment for Adolescents with Depression Study Team, 2004). Treatment foci were (a) pharmacotherapy management (educating patients about medication, adjusting dosage and dosage schedules, inquiring and dealing with side effects) and (b) clinical management (assessment of the client’s functioning in major areas of life, brief supportive counseling, and limited provision of advice). Fluoxetine (Prozac) typically was taken in the morning. The initial dose of 10 mg/day was increased to 20 mg/d during Week 1 and to 40 mg/d by Weeks 2 to 12. The maximum dosage allowed was 60 to 80 mg/d ($M = 50.1$, $SD = 4.5$). During Weeks 12 to 14, dosage was reduced again to 20 mg/d for 53% of participants whose depression improved, in keeping with standard practice (American Psychiatric Association, 2000). The remaining participants continued their prescribed dose during treatment. A maintenance treatment of fluoxetine (Prozac) at 20 to 40 mg/d was provided during the 6-month follow-up.

Booster sessions. After treatment completion, all participants in the cognitive conditions were allowed three booster sessions during the subsequent 6 months. Booster sessions could be scheduled at any time, so long as they were at least 1 month apart. For participants assigned to the fluoxetine condition, booster sessions focused on pharmacotherapy (e.g., inquiring about side effects) and clinical (e.g., brief supportive counseling and limited advice giving) management. All participants were asked to not pursue other treatment for depression through the 6-month follow-up.
Results

Costs of CT, REBT, and Fluoxetine (Prozac)

Total cost for each treatment condition, and cost per patient in each condition, was calculated by:

1. measuring the amount of each resource used for a service (e.g., 20 hr of therapy for each patient),
2. finding the unit cost of each resource (e.g., $100 for each hour of therapy),
3. multiplying the amount of each resource used by the cost per unit of that resource (e.g., 20 hr \( \times \) $100/hr = $2,000 per each patient’s therapy), and
4. summing the costs of each resource to arrive at the total cost, as detailed in Yates ([1980, 1996; e.g., Fals-Stewart, Yates, & Klostermann, 2005]).

Romanian cost figures were translated into U.S. dollars according to purchasing power parity (Gosden & Torgerson, 2002; International Monetary Fund, 2006).

Table 1 lists resources in the far-left column, and, moving from left to right in this Resource × Procedure matrix (cf. Yates, 1999), the unit measure used for each resource, the unit cost for each resource, the specific amounts of each resource used in each treatment condition, and total costs for the total amount of each resource used for each treatment. For example, on average, a participant assigned in the CT condition attended a total of 20.41 hr of therapy of a maximum of 23 hr (20 regular and 3 booster sessions). The cost of 1 hr of therapy was estimated at $6.59 based on the hourly wage of therapists in Romania (a weighted mean based on the fact that on each psychotherapeutic condition, there were three clinical psychologists and one psychiatrist who could provide the therapy). Total costs of treatment are listed in the bottom rows of the three right-most columns.

Treatment resources included providers’ time devoted to (a) preparation, (b) treatment, and (c) assessment (assuming that BDIs would be administered as part of therapy, even if no research was being conducted). The value of provider time was calculated per hour from monthly gross wages, including taxes paid by employers. Cost of office space was calculated according to office size, the proportion of available hours that space was used for preparation, treatment, and assessment, and rental costs for local real estate (cf. Yates, 1980, 1999).

Materials and investment costs included manuals and the BDI license and answer sheets. Cost of fluoxetine (Prozac) was obtained from a retail distributor’s price list, being the best offer from three other retail price lists (a common procedure when purchasing goods from public money). Indirect costs such as administration and space used for activities required by the interventions, such as supervision, were calculated by multiplying the total of the aforementioned costs by 1.50, reflecting a conservative overhead rate of 50%. Client resources devoted to interventions included time spent in sessions and transportation to and from sessions, multiplied by gross average national income plus fringe benefits, plus transportation costs using public transportation (see Yates, 1980, 1999).

The least expensive treatment was CT (averaging $504.84 per client) and REBT (\( M = $518.55 \) per client), followed by fluoxetine (Prozac) (averaging $666.94 per client). These differences reflect the higher costs associated with medical as opposed to psychological treatments and treatment settings. Moreover, medication was not included in overhead calculations because clients rather than providers typically purchase and retain medications. Statistical analyses could not be performed on
these data because cost data were collected for the treatment condition as a whole rather than for the individual client. Future research should examine individual variability in costs within treatment conditions (Yates, 1996).

Effectiveness of CT, REBT, and Fluoxetine (Prozac)

The three treatments seemed similar in their abilities to reduce depression between pretest and posttest, and to maintain this reduction 6 months following completion of treatment. Mean BDI scores for each condition are listed in Table 2. Analysis of variance considering type of treatment as a between-participants factor and time (pretest, posttest, and 6-month follow-up) as the within-participants variable found that BDI decreased significantly following each of the three treatments, $F(1.33, 184.83) = 383.94, p < .01$, using a Greenhouse–Geisser correction to adjust for
violation of sphericity (Field, 2000; Sava, 2004). Post hoc Bonferroni procedures showed that both posttest and follow-up BDI scores were significantly lower than pretest BDI scores, but found no significant change between posttest and 6-month follow-up. Type of treatment did not affect the decrease in depressive symptoms, as shown by a nonsignificant Treatment × Time interaction and a nonsignificant effect for treatment type, \( F_{S} < 1 \). Similar findings were evidenced on the Hamilton Rating Scale for Depression, as detailed in David et al. (2008).

Cost-Effectiveness of CT Versus REBT Versus Fluoxetine (Prozac)

As detailed next, $/DFD was calculated for a client by (a) transforming the client’s BDI scores into DFDs, (b) calculating DFDs gained by the client after intervention, and (c) dividing DFDs gained by the cost of treating a client using the assigned treatment (from Table 1). Mean $/DFD gained per month are presented in the last two rows of Table 3.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>CT ((n = 49))</th>
<th>REBT ((n = 48))</th>
<th>Fluoxetine (Prozac) ((n = 44))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest scores</td>
<td>29.47, 8.95</td>
<td>32.73, 11.24</td>
<td>31.02, 12.09</td>
</tr>
<tr>
<td>Posttest scores</td>
<td>9.45, 6.72</td>
<td>9.54, 6.78</td>
<td>10.57, 6.51</td>
</tr>
<tr>
<td>6-month follow-up scores</td>
<td>9.63, 6.65</td>
<td>9.19, 6.22</td>
<td>10.25, 5.93</td>
</tr>
</tbody>
</table>

Table 2

Means and SDs of Beck Depression Inventory (BDI) Scores for Cognitive Therapy (CT), Rational-Emotive Behavior Therapy (REBT), and Fluoxetine (Prozac)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>CT ((n = 49))</th>
<th>REBT ((n = 48))</th>
<th>Fluoxetine (Prozac) ((n = 44))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest scores</td>
<td>29.47, 8.95</td>
<td>32.73, 11.24</td>
<td>31.02, 12.09</td>
</tr>
<tr>
<td>Posttest scores</td>
<td>9.45, 6.72</td>
<td>9.54, 6.78</td>
<td>10.57, 6.51</td>
</tr>
<tr>
<td>6-month follow-up scores</td>
<td>9.63, 6.65</td>
<td>9.19, 6.22</td>
<td>10.25, 5.93</td>
</tr>
</tbody>
</table>

Table 3

Means and SDs for Depression-Free Days (DFDs) and Gain in DFDs per Month

<table>
<thead>
<tr>
<th>Treatment</th>
<th>CT ((n = 49))</th>
<th>REBT ((n = 48))</th>
<th>Fluoxetine (Prozac) ((n = 44))</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFDs/month before treatment</td>
<td>56, 5.06</td>
<td>57, 3.92</td>
<td>57, 5.24</td>
</tr>
<tr>
<td>DFDs/month after treatment</td>
<td>55, 23.72</td>
<td>57, 23.76</td>
<td>52, 23.30</td>
</tr>
<tr>
<td>DFDs/month follow-up</td>
<td>49, 24.46</td>
<td>48, 25.08</td>
<td>47, 23.70</td>
</tr>
<tr>
<td>DFDs posttest–pretest (gain)</td>
<td>55, 18.57</td>
<td>57, 19.83</td>
<td>49, 18.10</td>
</tr>
<tr>
<td>DFDs follow-up–pretest (gain)</td>
<td>49, 19.23</td>
<td>48, 21.31</td>
<td>47, 18.39</td>
</tr>
<tr>
<td>$ per DFDs gained per month, posttest–pretest</td>
<td>55, 70.76</td>
<td>57, 56.46</td>
<td>49, 91.14</td>
</tr>
<tr>
<td>$ per DFDs gained per month, follow-up–pretest</td>
<td>49, 44.40</td>
<td>48, 29.85</td>
<td>47, 80.81</td>
</tr>
</tbody>
</table>

Transforming BDIs into DFDs. Using David’s (2006) norms, BDI \( \leq 7 \) indicated absence of depression, and a BDI \( \geq 29 \) indicated severe depression. Thus, if a client scored 7 or lower, the client was assumed to have a DFD every day for the month (i.e., 30 days) preceding the assessment and was given a DFD weighting of 1. Similarly, if a client had a BDI score of 29 or higher, the client was assumed to have had no DFDs during 30 days preceding administration and was given a DFD weighting of zero. Clients with BDIs between 8 and 28 were assigned DFD weightings found by linear interpolation. For example, a BDI of 10 was assigned a DFD weighting of \( \frac{29 - 10}{29 - 7} = \frac{19}{22} = .86 \). DFDs were calculated for the
client by multiplying the client’s DFD weighting by the period (30 days, for the present study) preceding assessment (e.g., \( .86 \times 30 = 25.8 \) DFDs per month).

**DFDs gained per month.** DFDs gained per month were found by subtracting pretest DFDs from posttest DFDs, and by subtracting pretest DFDs from follow-up DFDs. For example, if a client experienced 3 DFDs during the month preceding pretest, and 8 DFDs during the month preceding the posttest, the client gained 5 DFDs for the month. Means and SDs for DFDs gained are presented in Table 3. The overall distribution of DFDs was reasonably symmetrical, as suggested by a histogram and skewness of \(-.21\) and kurtosis of \(-1.45\). These were not reasonably similar to the distribution characteristics found for BDIs overall, with skewness of \(.92\) and kurtosis of \(.40\). As expected from the findings for BDI, no significant differences were found between treatments in DFDs gained.

**Cost per DFD gained per month.** For each client, total cost per client in a specific treatment (Table 1) was divided by DFDs gained for a client to calculate that client’s \$/DFD (e.g., \$504/18.32 \text{DFDs} = \$27.51 \text{per DFD gained per month, pretest vs. posttest}). To avoid division by zero and negative numbers, \$/DFD for participants with no DFD gain (five clients) or DFD decreases (one client) were set as the cost of treatment for that participant. Possible differences in \$/DFD gained per month between treatments were examined with Kruskal–Wallis ANOVAs rather than with parametric ANOVAs because log transformations failed to reduce skewness and heterogeneity of variance (Field, 2000; Sava, 2004). Significant differences in cost-effectiveness were found between treatments for both \$/DFD gained per month between pre- and posttest, \( \chi^2(2) = 15.59, p < .01 \), and between pretest and follow-up, \( \chi^2(2) = 18.77, p < .01 \). Post hoc analyses using mean ranks (Hinkle, Wiersma, & Jurs, 1994) revealed that both CT and REBT were significantly more cost-effective than fluoxetine (Prozac), in \$/DFD gained per month, and no significant differences were found between CT and REBT in \$/DFD gained per month, according to the same procedures.

**Cost-Utility Analysis**

Costs per QALY and QALY gained per year were calculated as detailed by Lave et al. (1998) using utility weights of \(.40\) for severe depression and \(.90\) for recovery from depression (Bennett, Torrance, Boyle, & Guscott, 2000; Freed, Rohan, & Yates, 2007; Lave et al., 1998; Revicki, Brown, & Palmer, 1995; Revicki & Wood, 1998) (see Figure 1). Treatments’ \$/QALY and \$/QALYG were compared by Kruskal–Wallis ANOVAs for the same reasons presented earlier. Treatments differed significantly both in \$/QALY, \( \chi^2(2) = 35.79, p < .01 \), and in \$/QALYG, \( \chi^2(2) = 18.66, p < .01 \). Post hoc analyses using mean ranks (Hinkle et al., 1994) found that both CT and REBT had better more cost-utility (i.e., had lower \$/QALY and lower \$/QALYG) than did fluoxetine (Prozac). No significant differences were found in \$/QALY or \$/QALYG between CT and REBT. Median \$/QALY was \$1,638 for CT; \$1,734 for REBT; and \$2,287 for fluoxetine (Prozac). Median \$/QALYG was \$2,120 for REBT; \$2,342 for CT; and \$3,162 for fluoxetine (Prozac).

**Discussion**

The present study endeavored to measure and statistically contrast the effectiveness, costs, cost-effectiveness, and cost-utility of CT, REBT, and fluoxetine (Prozac) in treating MDD using a randomized, controlled clinical trial design in a setting
that could yield findings more generalizable to more countries and more people than findings reported to date. In what we have argued to be the more internationally representative context of an upper middle income economy, two forms of CBT, and pharmacotherapy, proved to be effective—equally effective—in reducing depression according to measures advocated by both researchers (BDIs) and policy makers (DFDs and QALYs). CBT even showed some advantage, on some measures, over medication—as found in a variety of other studies (Butler et al., 2006). Both CT and REBT were substantially less expensive than was fluoxetine (Prozac) pharmacotherapy, a common drug prescribed throughout the world for depression. We replicated the effectiveness of evidence-based psychotherapies in treating depression (e.g., Antonuccio et al., 1995; Revicki et al., 2005; Schulberg et al., 2002) and demonstrated that important differences can be found between alternative treatments in their cost-effectiveness and cost-utility. According to median $/QALY, the same effect obtained by paying $100 for fluoxetine can be obtained by spending approximately $70 for CBT psychotherapy [71% less for CT vs. fluoxetine (Prozac), 75% less for REBT vs. fluoxetine (Prozac)].

Cost findings may, of course, be more sensitive than effectiveness findings to differences in economic context and to cost assessment methods. For instance, lower salaries for therapists providing psychological interventions in some countries could tilt cost even more in favor of CBT while higher salaries for therapists or lower costs for medication could tilt cost-effectiveness and cost-utility findings toward medication. Certainly, if trained therapists were unavailable but physicians skilled in monitoring dosage and possible side effects of antidepressant medication were, cost could well favor medication over CBT. In addition, shorter or less frequent sessions with psychiatrists would reduce costs of fluoxetine treatment, possibly markedly, as could replacement of psychiatrists by nurses or other drug-monitoring professionals. The latter was not allowed in the country in which this study was conducted.

Figure 1. Costs per QALY and QALY.
Additionally, new drugs could become less expensive to administer than fluoxetine (Prozac), making pharmacotherapy as or more inexpensive than CBT, although medical interventions for side effects of any such drugs would be important to include in cost calculations. For example, although Vos, Corry, Haby, Carter, and Andrews (2005) found lower cost per disability-adjusted life years (<AUS$ 10,000) using CBT and bibliotherapy, compared to SSRIs in Australia (AUS$17,000–20,000), they did find similar costs for CBT and TCA medication (which, as noted earlier, has more side effects than SSRIs; Antonuccio et al., 1997). Thus, the question may not be whether one form of therapy is more cost-effective, or has more cost-utility, than another, but to determine the economic and other circumstances in which one therapy is more cost-effective than others.

Cost findings also can be affected by the costing perspective adopted (e.g., provider vs. societal vs. consumer), the type of therapist used (e.g., doctoral vs. master’s vs. paraprofessional), the public or private setting in which therapy is provided, the duration of intervention and maintenance periods, and the overhead rate applied. Of particular interest is that our finding was obtained using a conservative rate of 50% for overhead costs. The lower the rate, the more competitive CBT psychotherapy becomes because the cost of medication was not included in the overhead. Therefore, reporting details of cost-measurement procedures and itemizing lists of the types and amounts of resources used rather than just reporting a single “bottom line” should facilitate comparisons between studies as well as better understanding of why different researchers arrive at different cost findings. Measuring costs for each client individually, as done by Fals-Stewart et al. (2005), might show more clearly the relationship between the resources invested in treatment and the outcomes produced for each client as well as better capturing the variability of costs over clients within the same treatment.

The present study introduces the use of DFDs and QALYs to the depression research literature, but these findings should be regarded as preliminary and illustrative. Our calculations of DFDs and QALYs assumed that BDI scores provided stable estimates of participants’ affect during the rather long periods between BDI administrations. Given that MDD can remit, recur, and otherwise change suddenly, it can be argued that BDIs provide only a point estimate of what can be a highly variable affective state. High variability of affect challenges our assumption that all points in time between BDI administration, including each DFD and each quality-adjusted time period, are sufficiently similar to infer, for example, that all days and time periods between end-of-intervention and follow-up BDIs are similarly free of depression. We hope that future research will assess depression in a manner that allows more definitive conclusions to be drawn regarding outcome measures that have gained great credence in other research literatures (cf. Gold et al., 1996).

The present study is in keeping with long-voiced standards for developing evidence-based treatments. As advocated by Chambless and Hollon (1988), “… in evaluating the benefits of a given treatment, the greatest weight should be given to efficacy trials but that these trials should be followed by research on effectiveness in clinical settings and with various populations and by cost-effectiveness research.” (p. 7). Here, we have reported findings for three classes of variables that provide crucial evidence regarding CBTs as well as a standard pharmacotherapy for depression: costs, effectiveness (both in terms of depression and QALYs), and cost-effectiveness (both cost of reducing depression and cost of adding QALYs).
In future research, we hope to collect more cost data directly from participants while they are receiving treatment instead of estimating costs. Follow-ups of 1 to 2 years rather than a half-year following completion of treatment could allow further differences in costs, and in effectiveness, to emerge. Although ethical issues in not treating persons with MDD argue against it, measurement only and attention placebo control conditions could allow more precise calculation of gains in DFD and QALY: To the extent that individuals with MDD spontaneously remit, using pretest scores may exaggerate gain in DFD and QALY that can be attributed to a specific treatment; however, note that this limitation would not affect costs or absolute values for total DFD or QALY. To persuade policy makers to adopt evidence-based psychotherapies such as CBT as primary treatments for MDD may require more than a few studies demonstrating the cost-effectiveness and cost-utility of those therapies (Haby et al., 2004). Including measurement of costs and analyses of cost-effectiveness and cost-utility are initial, but necessary, steps in providing more consumers with treatments that will benefit them the most with the least impact on limited healthcare resources.

References


