This is a chapter excerpt from Guilford Publications.
What Works for Whom?, Second Edition: A Critical Review of Psychotherapy Research, Anthony Roth and Peter Fonagy
Copyright © 2005

CHAPTER 4

DEPRESSION

DEFINITIONS

DSM-IV-TR describes a number of subcategories of depression; those particularly relevant to research studies are defined as follows (adapted from Wells, 1985).

Major Depressive Disorder

Major depressive disorder (MDD) is characterized by one or more major depressive episodes and the absence of manic episodes. A major depressive episode is defined by depressive mood or loss of interest or pleasure in almost all usual activities, accompanied by other depressive symptoms. These include disturbances in appetite, weight, and sleep; psychomotor agitation or retardation; decreased energy; feelings of worthlessness or guilt; difficulty concentrating or thinking; and thoughts of death or suicide, or suicidal attempts. DSM-IV-TR specifies that at least five of nine specific depressive symptoms must be present nearly every day for at least 2 weeks to make a diagnosis of MDD, and that the symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. Depressive episodes are distinguished from normal bereavement reactions.

Dysthymic Disorder

This disorder is characterized by depressed mood or loss of interest in nearly all usual activities, though symptom severity is not sufficient to meet the criteria for MDD. The disorder is, by definition, chronic. Symptoms should be present for at least 2 years, and a diagnosis cannot be made if patients are

symptom-free for more than 2 months in any 2-year period. It is characterized by depressed mood for most of the day, together with at least two of the following six symptoms: poor appetite, insomnia or hypersomnia, low energy, low self-esteem, poor concentration, and feelings of hopelessness. For diagnostic purposes, these symptoms should be severe enough to cause clinically significant distress or impairment in social, occupational, or other areas of functioning.

"Double Depression"

Patients with dysthymic disorder frequently present with a superimposed MDD; this is usually referred to as "double depression."

PREVALENCE AND NATURAL HISTORY

Prevalence

Only a portion of individuals with mental health problems present to family physicians or mental health professionals (e.g., Bebbington et al., 2000a; Goldberg & Huxley, 1980). Because of this, estimating treatment need is better done through community-based surveys rather than relying on data from clinical services. Two large-scale community surveys provide data on the prevalence of psychiatric disorders in the United States. The National Institute of Mental Health (NIMH) Epidemiologic Catchment Area (ECA) program was a five-site project sampling approximately 20,000 adults (Robins & Regier, 1991). The National Comorbidity Survey (NCS; Blazer et al., 1994) had a slightly more restricted age range, and interviewed approximately 8,000 adults between ages 15 and 54. The prevalence rates derived from these surveys need to be interpreted cautiously; for example, there is a risk that they are inflated by individuals whose distress is transient. Deriving a "correct" prevalence rate that accounts for the clinical significance of symptoms is difficult and controversial. Narrow et al. (2002) have recomputed prevalence rates from the ECA and NCS surveys, taking into account the degree to which symptoms resulted in help-seeking behavior and led to significant levels of distress (see Table 4.1). In addition, they attempted to reconcile differences in prevalence rates between the surveys, some of which relate to methodological differences. Their approach has been criticized as inappropriately robust (Wakefield & Spitzer, 2002), and it is clear, that while presentation to services appears to be linked to the severity of symptoms (Bebbington et al., 2000b), lack of help seeking cannot be assumed to indicate that distress is unimportant. Nonetheless, the revised rates are cited here (and in Appendix III), since they probably yield a more accurate indicator of service need.

TABLE 4.1. One-Year Prevalence Rates for MDD and Dysthymic Disorder for Adults Ages 18–54

	ECA (corrected for clinical significance)	NCS (corrected for clinical significance)	Combined ECA + NCS estimate after correction for clinical significance
MDD	5.4 (4.6)	8.9 (5.4)	4.5
Dsythymia	5.7 (1.7)	2.5 (1.8)	1.6

Note. Data from Narrow et al. (2002).

The ECA and NCS estimate 1-year prevalence for MDD at 5.4% and 8.9%; corrected for clinical significance, these figures lower to 4.6% and 5.4%, respectively. Narrow et al.'s (2002) estimate, which combines data from both surveys, is 4.5%. For dysthymic disorder, the ECA and NCS estimates are 5.7% and 2.5%, respectively; with correction for clinical significance, these reduce to 1.7% and 1.8%, respectively; the combined estimate is 1.6%. Other reviews derive somewhat different (uncorrected) rates. Angst (1992), reviewing 17 studies, suggests that 1-year prevalence rates for MDD lie between 2.6 and 6.2%, and for dysthymic disorder, between 2.3 and 3.7%. Lifetime prevalence rates vary between 4.4 and 19.5%. Angst also reports data from a Swiss prospective community survey carried out (to date) over 10 years. This was based on multiple interviews and hence avoided problems of estimating prevalence based on recall. Lifetime prevalence to age 30 of MDD was 14.5%, with around half of affected individuals seeking treatment.

The prevalence of depression varies by gender and age; prevalence of MDD in the ECA and NCS was almost twice as high in women as men, and greater in younger adults. In part, this may reflect the greater willingness of younger adults to admit to mental health problems (Taube & Barrett, 1985; Weissman et al., 1988), or problems of recall when older respondents are interviewed in cross-sectional surveys (Fombonne, 1994). However, there is evidence that prevalence within younger age groups is increasing (Burke et al., 1991), though the degree to which this is associated with comorbid substance abuse is unclear. Furthermore, there is some agreement that overall rates of depression are increasing (Fombonne, 1998b; Klerman & Weissman, 1989).

Natural History

Most studies of "natural" history monitor longitudinal outcomes for patients offered "treatment as usual" (TAU). Over a 2-year period, Wells et al. (1992)

followed up 626 outpatients; the sample included patients diagnosed with MDD, dysthymic disorder, and double depression, and also contained clients with subthreshold depressive symptoms. Patients with MDD had a 42% probability of remission[1] in the first year, and a 60% probability of remission in the second year, if none had occurred in the first year. Clients with double depression had a rather different course, depending on the severity of their symptoms. Those with more severe symptoms had a 37% likelihood of remission in the first year; if no remission occurred by this point, there was only a 16% probability of remission in the second year. Both dysthymic patients and those with subthreshold symptoms of depression were at considerable risk of suffering an episode of MDD over the study period. Half the patients with an initial diagnosis of dysthymic disorder and 25% of patients with subthreshold symptoms of depression (with or without a prior history of depression) experienced an episode of MDD over the 2-year period. Data from patient samples in field trials for DSM-IV-TR confirm this pattern; 79% of patients with dysthymic disorder eventually developed MDD (McCullough et al., 1992). The poorest clinical outcomes were found in patients with double depression; there was a particularly low rate of remission in patients with both double depression and high initial symptom severity. Patients with dysthymic disorder (even in the absence of MDD) had higher levels of depressive symptoms over the 2-year period of the study than patients with MDD alone, despite the fact that dysthymic disorder is defined by the presence of less severe (if persistent) depressive symptoms. In addition, patients with dysthymic disorder were rated as having poorer social and emotional functioning than those with MDD.

Keller and Shapiro (1982) and Keller et al. (1983) suggest that patients with double depression tend to have a shorter episode of MDD but are also likely to relapse more quickly than those with MDD alone. Double depressives appear to have a faster "cycle time"; over a 2-year period, 62% of them had completed a cycle of recovery and relapse, compared to 33% of the MDD group.

---

[1] Recovery and relapse are problematic terms unless specified, and are used inconsistently in the literature (Prien et al., 1991). Frank et al. (1991) have proposed the adoption of the following definitions, many of which have been adopted by researchers but (where relevant) with somewhat differing time courses:

Partial remission: A period when improvement of a sufficient magnitude is observed, but during which the patient continues to show more than minimal symptoms.

Full remission: A relatively brief period during which the individual is asymptomatic.

Recovery: A remission of a longer period, usually indicating recovery from the index episode (though not from the illness per se).

Relapse: A return of symptoms satisfying criteria for the full syndrome that occurs during the period of remission, but before recovery.

Recurrence: The appearance of a new episode of MDD arising during a period of recovery.

Long-term monitoring confirms a pattern of vulnerability to relapse for people with MDD. Piccinelli and Wilkinson (1994) reviewed 50 naturalistic follow-up studies of in- and outpatients with unipolar depression, carried out between 1970 and 1993. Although recovery rates increase over time (on average, 53% of patients will recover at least briefly by 6 months), one-fourth of the patients will have suffered a recurrence of the index episode within 1 year. Seventy-five percent of patients followed up for 10 years suffered a further episode of depression, and 10% of patients suffered persistent depression. Mueller et al. (1999) followed up patients over 15 years; all were in receipt of TAU. Of 380 patients who had recovered from an index episode of MDD, a cumulative proportion of 85% relapsed over this period. Of a further 105 patients who had recovered and remained well over 5 years, a cumulative proportion of 58% relapsed. Though there were indications that TAU included suboptimal delivery of medication, there was little information available regarding the use of psychosocial interventions. Demographic or clinical characteristics did not predict relapse, though there were indications that individuals who had recovered but continued to experience subthreshold symptoms were particularly vulnerable, a pattern found in other studies (e.g., Judd et al., 1998b).

Summary

Studies of the prevalence and natural history of depression have a number of implications for research. Although precise estimation is complicated, it is clear that depression is a relatively common syndrome affecting at least 4.5% of the population, with prevalence among women about double that among men. The course of depression appears to differ according to subtype (MDD, dysthymic disorder, or double depression). It is likely that 80% of patients with dysthymic disorder will eventually develop an MDD, suggesting that dysthymic disorder and acute depression are variants of a similar condition. Relapse is a serious challenge: 85% of patients followed up over 15 years, and 75% of patients followed up over 10 years will have suffered a further episode of MDD, and 10% of these will have endured persistent depression. The probability of relapse is increased in patients with more than three previous episodes of MDD but is greatest in patients with a diagnosis of dysthymic disorder; these patients show a faster cycle of recovery and relapse than those with MDD alone. Even among those patients who have "recovered," subthreshold symptoms are common and are associated with an increased likelihood of relapse.

The risk—indeed, the probability—of relapse has obvious implications for treatment trials. The effectiveness of a treatment needs to be judged not only by its capacity to manage an index episode but also by its ability to maintain remission. This poses a challenge, in part, because on the basis of

figures given above, long-term follow-up of at least 2 years would be necessary to provide a conclusive result that is not confounded with the natural history of this disorder. It is also likely that outcome in clinical trials will be influenced by case mix, and particularly by the presence of patients with double depression or a history of recurrent MDD. Because of the exclusion criteria applied in at least some research trials, it is possible that the clinical population will contain comparatively more patients with chronic depression and dysthymic disorder. This may lead to overestimation of treatment effects; poorer outcomes might be expected in clinical practice than in trials. However, as an increasing number of studies concern themselves with "treatment resistant" patients, this may be a less pertinent issue than before.

LANDMARK STUDIES OF EFFICACY

Subsequent chapters review individual studies in the context of meta-analyses and qualitative reviews. This chapter adopts a different strategy, describing in some detail a small number of high-quality individual studies that help to contextualize the broader body of evidence. These trials give indications of the acute and longer term efficacy of the major treatment approaches in this area (cognitive-behavioral therapy, interpersonal psychotherapy, short-term psychodynamic therapy, and medication), and of the challenge posed by relapse.

Cognitive-Behavioral Therapy and Interpersonal Psychotherapy: NIMH Treatment of Depression Collaborative Research Program

This major and widely cited research program (summarized in Elkin, 1994) set a standard against which other studies can be judged. Patients were randomized to receive one of four interventions: cognitive-behavioral therapy (CBT; Beck et al., 1979), interpersonal psychotherapy (IPT; Klerman et al., 1984), imipramine plus clinical management (IMI-CM), or placebo plus clinical management (PLA-CM). Clinical management consisted of a weekly meeting of 20–30 minutes to discuss medication, side effects, and the patient's clinical status. In addition, and where necessary, support, encouragement, and direct advice were also offered. On this basis, it is worth noting that both medication conditions contained psychotherapeutic elements. This research design has been misinterpreted as a test of therapy against medication (Elkin, 1994); more accurately, the intention was to use medication condition as a "benchmark" against which to compare the psychological therapies.

The study was carried out at three research sites in the United States. Five hundred sixty outpatients were initially screened, essentially ensuring that patients met criteria for a DSM-III-R diagnosis of unipolar depression.

Two hundred fifty patients, all moderately to severely depressed, were selected for the trial; 239 actually entered it. Of these, 60% had been depressed for more than 6 months; for only 36% was this a first episode of depression.

Treatments were carried out by experienced therapists (10 each in IPT and pharmacotherapy, eight in CBT) chosen for their expertise in applying their respective therapy and supervised regularly throughout the clinical trial. To ensure that therapies were conducted as intended, sessions were taped, and the process of therapy was checked against measures of therapy adherence. Though there were some differences in attrition from each condition, these were not statistically significant. Rates of dropout across treatment modalities were as follows: 23% (n = 14) for IPT, 32% (n = 19) for CBT, 33% (n = 19) for IMI-CM, and 40% (n = 25) for PLA-CM.

Patients were assessed before treatment and at 4, 8, 12, and 16 weeks, and followed up at 6, 12, and 18 months. During therapy, a number of standardized measures of symptomatic status were employed (including the Hamilton Rating Scale for Depression [HRSD] and the Beck Depression Inventory [BDI]). After discharge, progress was assessed using a semistructured interview designed to assess the longitudinal course of psychiatric disorders (the Longitudinal Interval Follow-Up Evaluation II [LIFE-II]; Keller et al., 1987).

Analyses were carried out on three overlapping sets of patient samples: a completer sample (n = 155) that had received at least 12 sessions or 15 weeks of therapy; the sample of patients that had entered treatment and received at least four sessions of therapy (n = 204); and the total (intent to treat) sample of patients that had entered the trials (n = 239).

Posttherapy, the general direction of results was similar on all measures and in all samples (Elkin et al., 1989), with patients who received IMI-CM having the lowest symptomatic scores, PLA-CM the most symptomatic, and the psychotherapies in between and usually closer to IMI-CM. The magnitude of these differences was not large, and pairwise comparison of treatment conditions revealed no differences between therapies or between therapies and IMI-CM.

In addition to comparisons of relative scores, a comparison of "recovery rates" was carried out—clearly a more stringent and clinically relevant analysis. Recovery was defined as an HRSD score ≤ 6 and a BDI score ≤ 9. No significant differences between treatment groups were found employing the BDI data, though, in part, this seems to reflect the degree of improvement in the PLA-CM condition. Using HRSD data, significant differences were apparent. Pairwise comparisons using the complete sample of patients indicated that those who received IMI-CM and IPT were significantly more likely to recover than those who had received PLA-CM; a trend toward sig-

nificance was apparent in the other two patient samples. There were, however, no significant differences between therapies or between therapies and IMI-CM in any patient sample.

All treatment conditions resulted in a significant improvement pre- to posttreatment—perhaps surprisingly, this included PLA-CM. Outcome for patients in the therapy conditions was equivalent to that in other treatment trials (reviewed below). The lack of significant differences seems attributable to the good performance of PLA-CM. However, as already noted, PLA-CM does contain a number of nonspecific therapeutic elements; in a sense, comparisons between it and the psychotherapies may reflect differences between such elements and the technical interventions embodied in the therapies.

Ogles et al. (1995) reanalyzed data for the completer sample of patients, arguing that the Elkin et al. (1989) analysis did not consider the reliability of change from pre- to posttherapy, that there was no attempt to consider the clinical significance of multiple measures simultaneously, and that the possibility of reliable deterioration was not determined. Using Jacobson and Truax's (1991) method, it is apparent that the proportion of clients showing reliable change was statistically equivalent across treatments, with little evidence of deterioration—in fact, observed at a rate of between 2 and 5% dependent on the measure employed. If clinically significant change is determined by movement into the functional distribution of scores across instruments (within 2 standard deviations of the mean for the normal population), the immediate impact of therapy across all treatments was equivalent when using the BDI and the HRSD. This further emphasizes the substantial improvements achieved by clients within the PLA-CM group (e.g., 62% achieved clinically significant change using the BDI as a measure). However, there were significant differences across treatments on measures of general symptomatology (using the Hopkins Symptom Checklist [HSCL-90, Lipman et al., 1979], equivalent in most respects to the Symptom Checklist 90 [SCL-90; Derogatis, 1977]). Thus, on this instrument, 78%, 93%, and 87% of the CBT, IPT, and IMI-CM groups, respectively, were placed in the functional distribution, contrasted to 65% of the PLA-CM group.

Although (as noted earlier) Elkin et al.'s (1989) analysis found statistically significant differences between treatments only when the HRSD was employed as a measure, Ogles et al. (1995) noted that reanalysis for both clinical significance and concordance across measures showed a high level of agreement. Thus, for 118 (73%) of 162 clients, all three measures were in agreement regarding the clinical significance of the change pre- to posttreatment.

A secondary analysis (Elkin et al., 1989) considered the degree to which initial symptom status influenced outcome. Two definitions of severity were employed. In the first, severity was defined as an HRSD score of ≥ 20; in the second, a Global Assessment Scale (GAS) score of ≤ 50 was used (tapping

both depressive symptomatology and functional impairment). The results suggested that for the less depressed group, there were no significant differences between treatments. However, for the more depressed group (and contrasting the three patient samples), pairwise comparisons using the HRSD revealed consistently lower scores for IMI-CM than for PLA-CM, and some significant differences for IPT compared to PLA-CM. Using the recovery criterion, there were no significant differences between treatment groups for less severely depressed patients, but, again, patients receiving both IMI-CM and IPT were significantly more likely to recover than those in PLA-CM. However, no significant differences were found between therapies.

Reanalysis of these data (Elkin et al., 1995) using more powerful statistical techniques (random regression models) confirms the equivalence of intervention methods for less depressed patients but indicates greater differentiation among the therapies for the more depressed sample. For these patients, using both HRSD and BDI scores as outcome measures, IMI-CM and IPT appeared equally effective. IMI-CM was significantly more effective than CBT or PLA-CM; IPT showed a trend ($p < .08$) toward greater efficacy. CBT was no more effective than PLA-CM. When the GAS was used as an outcome measure, a different pattern emerged, with IMI-CM being more effective than the other three interventions, all of which showed equal relative efficacy.

These analyses are exploratory, but they do suggest that initial patient severity may be an important factor in considering treatment allocation—particularly the finding that for patients with lower levels of depression, PLA-CM (which could be considered a "minimal support" intervention) was as effective as the active therapies. Although there is evidence that IMI-CM was particularly effective with more severely depressed and functionally impaired patients, it should be noted that it was no more effective than IPT when patients were symptomatically rather than functionally impaired. A number of researchers have interpreted these results to indicate that medication is necessarily the treatment of choice in more severe depression (Elkin, 1994). These results suggest that some caution should be taken in this regard.

Follow-up of patients continued over 18 months and is reported in Shea et al. (1992a). In this analysis, the question of interest was the fate of the patients who had met a stringent criteria for recovery—at least 8 weeks following completion of treatment with minimal or no symptoms. Relapse was defined as the presence of at least 2 weeks of MDD-level symptoms over the 18-month follow-up period.

Only 20% of the original sample and 24% of the patients with follow-up data met the criteria for recovery with no relapse. Of those entering therapies, 24% of those receiving CBT remained recovered without relapse at 18 months, compared with 23% for IPT, 16% for IMI-CM, and 16% for PLA-

CM. Of those with follow-up data at 6 months, 30% in CBT, 26% in IPT, 19% in IMI-CM, and 20% in PLA-CM were recovered without relapse at 18 months. Despite the presence of a trend for psychotherapy to be superior to IMI-CM, there are no statistical differences in these rates. However, it should be noted that, as is the case for any study that attempts a naturalistic long-term follow-up, there are problems in interpreting the data statistically: Because the groups no longer benefit from the original randomization, attributions about effects cannot be traced to the treatment modality employed (Shea et al., 1992a).

Overall, it is clear that rather few patients recover and remain well with 16 weeks of treatment, and a clear conclusion from this study is that interventions of this length are not sufficient to maintain functioning in the majority of patients. This result is not, perhaps, surprising in the light of evidence considered earlier regarding the natural course of depression.

Although the differences between CBT and IPT in this trial were not substantive, relative to other studies, CBT did perform less robustly than many of its practitioners would have expected, and this has led to debate about the competence of its delivery and supervision (e.g., Hollon & Beck, 1994). There are suggestions that CBT performed much better at one site, and IPT performed more poorly at another (Elkin, 1994). Though these data have not been confirmed in later analyses (Elkin et al., 1996), these results have continued to be contested (Jacobson & Hollon, 1996a, 1996b). Unfortunately, even in a trial as large as this, statistical power reduces markedly when investigators seek site differences for specific forms of therapy conducted with high-severity patients, because at this level of stratification, the number of available "data points" is very low.

Notwithstanding the need to accept the data as reported, there may be some indications that CBT and IPT were delivered at different levels of competence, though whether this reflects an actual or methodological difference is unclear. As before, sessions were taped and monitored, and therapists were alerted when their performance was felt to be problematic. On this basis, alerts were issued for 33% of monitored CBT tapes but only 3% of IPT tapes (Elkin, 1999). However, CBT and IPT supervisors had different procedures for scoring competence, and in some respects, the standards for CBT do seem to have been more stringent. While this may—or may not—be grounds for questioning the "adequacy" of treatment delivery, it may also raise questions about the impact of deviations from treatment fidelity in routine clinical settings, where expert therapists and intensive supervision are rarely available.

The data set from this trial is now publicly available, and a number of post hoc analyses have been undertaken, attempting to link therapy, therapist, and patient characteristics to outcome. These are considered later in this chapter and in Chapter 16, along with other process studies.

Cognitive-Behavioral Therapy Alone and in Combination with Medication: University of Minnesota and University of Pennsylvania–Vanderbilt University Studies

Perhaps because of its scale and rigor, findings from the NIMH trial have been highly influential, particularly in its conclusions regarding the relative efficacy of CBT and medication. This issue is addressed in two trials by Hollon, DeRubeis, and colleagues, who have examined the impact of these therapies in moderately to severely depressed individuals.

The first trial at the University of Minnesota (Evans et al., 1992; Hollon et al., 1992) aimed to examine both response to acute treatment and patterns of posttherapy relapse. On this basis, 107 patients (all with BDI scores ≥ 20) were randomized to one of four treatment arms. In the first, three patients received 12 weeks of acute treatment: CBT alone, IMI-CM, or CBT and imipramine combined. The fourth condition comprised 12 weeks of acute treatment with imipramine, followed by 12 months of maintenance medication. Clinical management was similar to that employed in the NIMH trial. Treatments were administered by experienced therapists and were somewhat more intensive than in the NIMH study, with 20 sessions planned for the 12-week period. Sixty-four patients completed all treatments; although attrition was high, there was no significant difference in the rate between conditions.

At the end of treatment (and, at this stage, considering both medication conditions together), all three treatments showed equal efficacy, though with a nonsignificant trend toward better results for the combined treatment group. In contrast to the NIMH study, there was no indication of a differential response with more severely depressed patients (although the sample size was perhaps too small to allow this comparison).

At 2-year follow-up, clear differences were found between treatment groups. Recovery was defined along the same lines as in the NIMH trial, though relapse was indicated by a consistently elevated BDI score. Although 44 patients (of the 64 completing treatment) were followed up, the low number of patients in each treatment cell necessitates some caution in interpreting the treatment results. Nonetheless, adjusting for patient attrition from the study, patients receiving medication without continuation showed the greatest rate of relapse (50%), all within the first 4 months of follow-up. In contrast, the relapse rate of patients receiving cognitive therapy (either alone or in combination with medication) was 18%. Most of these relapses occurred later than in the medication–no–continuation condition; mean survival times were 17.4 ± 1.2 months and 3.3 ± 0.4 months, respectively. Relapse rate in the medication continuation condition was intermediate, with a 32% relapse rate and a mean survival time of 17.3 ± 2.1 months. A secondary analysis indicated that relapse rates in the two cognitive therapy conditions were not different from one another.

More recent (and more substantive) research conducted at the University of Pennsylvania and at Vanderbilt University extends this work in a larger sample of moderately to severely depressed patients treated both with CBT and selective serotonin reuptake inhibitors (SSRIs) (DeRubeis & Amsterdam, 2002; Hollon & Shelton, 2002). Entry criteria included a current episode of MDD, with an HRSD score ≥ 20. Two hundred forty patients met these criteria, and selection for appropriate severity and refractoriness appears to have been successful: Approximately half met criteria for chronic depression, with a mean duration of the current episode of 46 months; about 75% met criteria for recurrent depression, with a mean of 2.4 prior episodes of MDD.

In order to carry out the study of relapse described below, treatment assignment was asymmetric: 120 patients received antidepressant medication, 60 received pill placebo, and 60 received CBT. Active treatments were delivered over 16 weeks; those in receipt of placebo were monitored over 8 weeks, at which point active treatment was offered (on the basis that withholding active intervention over a longer period would be unethical). Medication was delivered by pharmacotherapists expert in treating depression, who met with patients once weekly in the first 4 weeks, then biweekly for the remaining 12 weeks. All patients were initially treated with paroxetine; those who showed no evidence of a response at 8 weeks had their medication augmented with lithium, desipramine, or Wellbutrin.

A similar pattern of results was obtained using both a categorical analysis (which identified patients as responders if they had an HRSD score ≤ 12) and hierarchical linear modeling (a form of growth curve analysis that models the progress of each individual patient rather than aggregating patients' scores in the form of means). At 8 weeks, medication patients showed a faster rate of response than those receiving CBT or placebo; respective response rates were 50%, 43%, and 25%. At 16 weeks, both active treatments showed an equivalent response rate of 58%, though there was a significant treatment-by-site interaction; on one site, CBT was superior to medication, at the other, medication was superior to CBT. Response rates at each site for medication were 48% and 67%, and for CBT, 63% and 53%. Further post hoc analysis is required to determine any reasons for these differences, but initial observation suggests that on one site, therapists were less experienced in CBT, and their patients tended to be more comorbid (two factors that might be expected to reduce response rates). Cross-site differences in CBT were less striking than for medication. While differences in prescribing practice may be relevant, it is not clear why such a large difference should emerge, especially given the particular care taken in this study to ensure that pharmacotherapy was delivered according to the highest standards.

One hundred four patients (58% of the sample assigned to active treatment) were classified as treatment responders and entered the next 12-month follow-up phase of the study, which focused on relapse prevention. The 69

responders to medication were randomized either to continue their medication or to be switched to pill placebo; all continued to receive regular clinical management. Patients who had responded to CBT (n = 34) were permitted up to three maintenance sessions. Relapse was defined by an HRSD score ≥ 14 over 2 successive weeks, or if patients met criteria for MDD. Patients who had received CBT were significantly less likely to relapse than those switched to placebo (31% vs. 76%); patients maintained on medication showed statistically equivalent relapse rates to those who had received CBT (47%). Adjusting these figures for patients who were nonadherent to medication gave a relapse rate for continuation medication of 42%.

Taken together, these studies suggest that CBT may be a robust intervention in individuals with moderate to severe depression. The more recent trial was conducted to rigorous methodological standards, and indicates that the capacity of CBT both to achieve and to maintain recovery from acute episodes is equivalent to a regimen of carefully administered maintenance pharmacotherapy, This result stands in contrast to that from the acute-phase of the NIMH study, though the extent of cross-site differences—most particularly for medication—reinforces the need for caution in interpreting outcomes from individual studies, no matter how well conducted.

Cognitive-Behavioral Therapy and Short-Term Psychodynamic–Interpersonal Psychotherapy: Sheffield Psychotherapy Project

As well as contrasting CBT and short-term psychodynamic/interpersonal therapy (and representing a rare empirical test of the latter technique), this trial was designed to explore a number of methodological and clinical issues raised by prior research. Most pertinent to this review, the study design stratified patients in relation to symptom severity and allocated them to treatments of different length (either 8- or 16-session treatments). The first issue emerged as a concern in the NIMH trial, but on the basis of a secondary analysis of their data. The second reflects early work on dose–response relationships (Howard et al., 1986), which suggested that treatment gains are most rapid early in therapy, with improvements in subsequent sessions showing a negatively accelerating "dose–response relationship."

The study was therefore concerned with (1) the efficacy of the two different therapies, (2) the influence of initial symptom severity, and (3) the impact of offering differing lengths of treatment (and relatedly, any evidence for a differential speed of action between the therapies). A total of 257 patients was assessed; 169 met criteria for DSM-III-R and Present State Examination (PSE; Wing et al., 1974) definitions of major depressive disorder. Thirty-nine percent were referred by their family physicians or mental health services; the remainder were self-referred. Clients were stratified into

those with low (BDI score 16–20), moderate (BDI score 21–26), or high (BDI score ≥ 27) levels of depression.

Following assessment, patients were randomly allocated to CBT or psychodynamic/interpersonal therapy, with treatment lasting either 8 or 16 weeks. Because randomization took place in the context of stratification by the severity of depression, the design had 12 "cells," and patient allocation continued until each cell had been occupied by 10 patients (see Table 4.2).

Overall, both therapies were found to be equally effective, to exert their effects with equal rapidity, and to have equivalent results for clients at all three levels of symptom severity. However, an interaction was found between initial symptom level and duration of therapy. Patients with mild or moderate depression did equally well with either 8 or 16 weeks of therapy. In contrast, those with severe depression showed significantly better outcomes when they received 16 weeks of therapy compared to those who received only 8 weeks.

One hundred three of the 117 completer patients were followed up at 1 year (Shapiro et al., 1995) and classified as recovered (asymptomatic at least 4 months, defined as a BDI score < 9), as having relapsed (a BDI score ≥ 15 during a period of remission from the previous episode but before meeting the criterion for recovery), or as having a recurrence (BDI score ≥ 15 after the criterion for recovery has been met). Of the 103 patients, 52% were treatment responders: 57% maintained their gains, 32% partially maintained gains, and 11% relapsed or had a recurrence. Thus, the proportion of all patients entering the trial and remaining asymptomatic from posttreatment to 1-year follow-up is 29% (a figure comparable to that found in the NIMH study).

No overall differences were found in outcome or maintenance of gains between CBT and psychodynamic/interpersonal therapy, nor was the interaction between initial symptom severity and duration of therapy maintained. However, there was an interaction between treatment type and duration, with those patients receiving eight sessions of psychodynamic/interpersonal therapy doing less well at 1 year on all measures. In addition, there was a nonsignificant trend toward better maintenance of gains with 16-session CBT, contrasted to the three other treatment combinations. Furthermore, there was some evidence that the patients who were more depressed initially

TABLE 4.2. Allocation of 120 Patients to Treatment Cells in Shapiro et al. (1994)

Level of depression	Treatment type	Length of treatment
High	CBT *or* psychodynamic/exploratory	8 *or* 16 sessions
Moderate	CBT *or* psychodynamic/exploratory	8 *or* 16 sessions
Low	CBT *or* psychodynamic/exploratory	8 *or* 16 sessions

tended not to maintain their gains, regardless of treatment modality or duration.

Shapiro et al. (1994) tentatively raise a number of questions about the degree to which their findings have direct implications for service delivery. Their posttherapy data suggest that patients with mild or moderate depression will gain no more from 16 sessions than they would from eight; only patients with severe depression would derive extra benefit from (and hence justify) longer therapy. However, follow-up data suggest a more cautious interpretation, since the pattern of maintenance of gains suggests that simple recovery is not an adequate measure of the efficacy of brief interventions. In particular, eight sessions of exploratory therapy appears to be too little, and there is some evidence favoring 16 sessions of CBT. Overall, poorer maintenance was evident in patients with greater levels of initial distress. This may caution against too marked a contraction of therapy contact time, particularly for more depressed patients.

A further concern regarding this study is the extent to which its results can be generalized from a research to a clinical sample. Although it is clear that patients met study criteria for depression, the majority (approximately 60%) were self-referred or referred through occupational sources, raising some questions about the comparability of these patients to the usual clinical population. This question was partially addressed by the Collaborative Psychotherapy Project (CPP; Barkham et al., 1996a), an explicit attempt to replicate the Sheffield project within a standard clinical context. This was carried out by colleagues of those involved in the Sheffield project, and, though (for practical reasons) smaller scale, used a similar methodology and research design. Thirty-six patients of low, medium, and high depression severity were allocated to CBT or psychodynamic IPT for 8 or 16 sessions (see Table 4.2).

Two main effects were found. First, CPP patients fared significantly better in therapies carried out over 16 sessions than in 8-session treatments. Second, while posttherapy gains made by patients in the CPP and the Sheffield project were similar, at 3-month follow-up, there was strong evidence that CPP patients were failing to maintain their gains. The severity × duration interaction found at posttherapy in Shapiro et al. (1994) was not replicated, though there is good reason to believe that the low statistical power of the CPP study may have contributed to this null finding (Barkham et al., 1996a).

Taken together, the Sheffield study and the CPP provide evidence that both therapy modalities under study are—broadly—equally effective and equally rapid in their initial response. However, some caution is necessary in considering "dose–response" relationships found in the Sheffield project, particularly where these appear to indicate that very brief periods of therapy may be effective at posttherapy. Longer periods of therapy appear to be associated with better longer term outcomes, particularly in the case of psychodynamic/

interpersonal therapy. In addition, more severely distressed individuals have a greater risk of relapse, and, while there is no clear indication as to why, there is some evidence from the CPP that clinic-sample patients may be more at risk of relapse than those usually found in research populations.

The Impact of Maintenance Therapies on Relapse: University of Pittsburgh Study

Though the forgoing trials suggest that short-term psychotherapies can successfully impact on depressive symptomatology, a clear (and fairly consistent) finding is that after 1 year, only about one-fourth of a treated sample will remain well. The issue of relapse is an important one that is explored further later in this chapter. Frank and colleagues' study is unusual in that its primary aim was to examine the impact of maintenance treatment offered after short-term intervention had been completed (Frank et al., 1989, 1990, 1991; Kupfer et al., 1992).

Two hundred thirty patients were selected for inclusion in the trial on the basis of a history of recurrent depression. All had experienced at least three previous episodes of depression, with the preceding episode occurring no more than 2½ years before the index episode (the mean number of episodes was 6.8, with a median of 4). All were selected on the basis of a clear DSM-III-R diagnosis of MDD in the absence of other Axis I disorders. Those with double depression were excluded, as were patients with severe Axis II disorders.

All patients received short-term treatment with imipramine and interpersonal psychotherapy. Psychotherapy sessions were scheduled weekly for 12 weeks, then biweekly for 8 weeks, and then monthly. At whatever point that patients achieved remission (defined as a HRSD score or a Raskin Depression Scale score ≤ 5 for 3 consecutive weeks), a further 17 weeks of treatment was offered, during which Hamilton and Raskin scores had to remain stable. At this point, a third evaluation was carried out, and the 128 patients who had reached the recovery criteria were assigned to one of five maintenance treatments for 3 years, or until the recurrence of depression; treatments were offered monthly and consisted of (1) medication clinic/clinical management and imipramine, (2) medication clinic/clinical management and placebo, (3) IPT and imipramine, (4) IPT and placebo, or (5) IPT alone (see Table 4.3). It should be noted that all patients had been receiving IPT; those assigned to medication conditions alone continued to see their original therapist, though the nature of their interaction changed from one of therapy to clinical management, along the lines of the NIMH study (Frank & Kupfer, 1994).

Unusual both in research studies and probably in clinical practice, imipramine continued to be prescribed at high levels (a mean of 207 mg

TABLE 4.3. Mean and Median Survival Times in Five Maintenance Conditions for 3 Years

Treatment condition	Mean ± *SD* survival time (in weeks)	Median survival time (in weeks)
Medication clinic and imipramine	124 ± 13	[a]
IPT and imipramine	131 ± 10	[a]
Medication clinic and placebo	45 ± 11	21
IPT	82 ± 13	54
IPT and placebo	74 ± 12	61

Note. Data from Frank et al. (1990).
[a] Since 50% of these subjects did not have a relapse, no median can be calculated.

daily). Attrition from the maintenance phase of the study was relatively low; only 22 (17%) of patients assigned to this phase failed to complete the 3-year protocol.

Results are reported at 3 years for the main trial (Frank et al., 1990) and at 5 years for a further group of patients maintained on imipramine or placebo alone (Kupfer et al., 1992). At 3 years, and contrasted with patients receiving placebo, medication or the combination of medication with IPT resulted in a significant reduction in the relapse rate ($p < .0001$). Maintenance therapy with IPT or IPT and placebo also resulted in a significant though less marked reduction in relapse ($p < .043$). There was no advantage to combination treatment over imipramine alone.

Overall, over 3 years, patients treated with imipramine had a 22.6% recurrence rate contrasted with 78.2% for those on placebo. Patients treated with IPT (with or without placebo) had a 44.2% recurrence rate over the same period. Further analysis suggests that when the quality of treatment delivered was high, relapse rates with IPT were equal to those achieved with imipramine. Frank et al. (1991) examined audiotapes of 38 of the 52 patients receiving maintenance IPT (either alone or in combination with placebo) and used rating scales of therapy adherence to determine the degree to which this therapy was implemented as intended. Therapies were defined as high quality if the patients received IPT above the median of adherence ratings, or low quality if delivered below the median. Results were striking; patients receiving high-quality therapy had a median survival time to relapse of approximately 2 years, while those receiving low-quality therapy had a median time of only 5 months.

The quality of therapy delivered was not a reflection of "good" or "bad" therapists, since individual therapists implemented IPT accurately with some patients, and with less success with others. Although requiring further study, accuracy of implementation seemed to reflect an interaction between patient and therapist factors.

Patients completing the 3-year protocol who had been receiving active medication (with or without maintenance IPT) were invited to continue a further 2-year randomized trial of active medication against placebo (Kupfer et al., 1992). Twenty patients entered this trial, either continuing to receive the high-dose imipramine regimen, or being transferred to placebo medication. Thirteen patients continued to receive monthly IPT, evenly split across placebo and medication conditions. Again, survival times were significantly greater for patients receiving active medication (99.4 ± 4.4 weeks) than for those assigned to placebo (54.0 ± 14.6 weeks, $p < .006$). Only one-third of patients receiving placebo survived the study period without relapse; 78% of placebo survivors were receiving continuation IPT. Only 11% of patients receiving neither medication nor IPT survived without experiencing a relapse.

Summary

The studies reviewed in this section meet the most stringent criteria of methodological rigor and indicate the efficacy of IPT and CBT; in the single trial in which there was a contrast of dynamic-IPT and CBT, the two modes of treatment were equivalent in their efficacy. This broad equivalence between outcomes from bona fide therapies is an important finding, suggesting, as it does, that depression may be responsive to a range of psychotherapeutic techniques.

Contrast of pharmacotherapy to psychotherapy in these trials is open to interpretation, because the pattern of outcomes is rather complex. In the NIMH trial, a differential response favoring medication was apparent for more severely depressed clients, particularly in the contrast of pharmacotherapy and CBT. However, in both this and the University of Pittsburgh study, the efficacy of IPT was more or less equivalent to medication, though, in the latter trial, this was true only when it was well implemented. The University of Minnesota and University of Pennsylvania–Vanderbilt trials counterbalance findings from the NIMH study, not only indicating an equivalence of action between CBT and medication but also demonstrating this equivalence in samples selected for depression severity. Based largely on the NIMH trial, some commentators have recommended the first-line use of medication in cases of severe depression. Whatever the clinical merits of this approach, this is unjustified by evidence from the four trials reviewed to this point.

Although clinically significant change was observed in as few as eight sessions for less severely depressed patients, at least 16 sessions were necessary for more severely depressed individuals. This suggests that there is usually little justification for considering very brief treatments. However, whatever the immediate benefits, the long-term effectiveness of short-term treatments is relatively poor for the majority of patients: between only one-third and one-

fourth of any sample can be expected to be in remission after 18 months. This highlights the critical issue of relapse. The obvious implication of outcomes from these trials is that many patients will need further treatment beyond an initial intervention if they are to maintain optimal levels of functioning and avoid further episodes of depression.

All of these trials delivered therapies at high levels of quality, and there is some evidence (both in these studies and others) that better outcomes were obtained when therapists were more adherent to the planned treatment, and when they delivered therapy more competently. In order to achieve this, all studies ensured that there was extensive monitoring and supervision for therapists, and when practice was found wanting, remedial action was taken. The degree to which supervision contributes to outcome is unclear and largely untested, but because monitoring of fidelity is often seen as reflecting the requirements of research protocols, its potential contribution to achieved outcomes is often overlooked. In this sense, it may be important to remember that outputs from therapies in research contexts represent supervised rather than independent practice, and that this will have implications when implementing research findings into routine clinical settings.

A final methodological point is that even in these very well-controlled trials, there were a number of unexplained sources of variance—an obvious example of which was large treatment-by-site interactions. Inconsistency in outcomes between and within studies cautions against overinterpretation based on single-research cohorts and reinforces the need for methodological standards that enable researchers to identify and to report such unexplained variance.

QUANTITATIVE REVIEWS OF TREATMENTS FOR DEPRESSION

Many quantitative reviews of treatments for depression examine the relative efficacy of psychological treatments, as well as their efficacy when contrasted with adjunctive medication, or with medication alone. Three early reviews focus on contrasts between psychotherapy and pharmacotherapy (Conte et al., 1986; Quality Assurance Project, 1983; Steinbruek et al., 1983); these reviews excluded research examining the benefits of psychotherapy alone. Dobson (1989) and Neitzel et al. (1987) restricted their analyses to trials using the BDI as an outcome measure. Robinson et al.'s (1990) review is more comprehensive, examining a wider range of studies of psychotherapy and pharmacotherapy using multiple outcome measures and more diverse forms of therapy. The U.S. Public Health Service's Agency for Health Care Policy and Research (AHCPR) also reviews psychopharmacological and psychotherapeutic treatments, but with the more specific aim of developing clinical practice guidelines for primary care and other health care practitioners

(Depression Guideline Panel, 1993a, 1993b, 1993c). There is a tendency for more recent reviews to be concerned with specific issues. Gaffan et al. (1995) focus on the impact of researcher allegiance on reported outcomes, Cuijpers (1997) on bibliotherapy (discussed below), Gloaguen et al. (1998) on studies of CBT, Westen and Morrison (2001) on manualized psychotherapies, McDermut et al. (2001) on group treatments, Leichsenring (2001) on a contrast of CBT and psychodynamic therapy, and Churchill et al. (2001) on contrasts of psychotherapies. Though not formal meta-analyses, three "mega-analyses" aggregate data from previously published trials and examine outcomes for psychotherapy alone and combined with pharmacotherapy (Casacalenda et al., 2002; Thase et al., 1997), and for CBT and pharmacotherapy (DeRubeis et al., 1999).

Aggregating information across these meta-analyses is complicated by the fact that each uses different inclusion and exclusion criteria, and sometimes employs different metrics for the analysis. In some cases, they do not include a full list of included studies (Neitzel et al., 1987; Quality Assurance Project, 1983; Steinbruek et al., 1983), making it hard to be clear about the degree of overlap between reviews (though this is indicated where feasible). Earlier meta-analyses also inevitably include a number of studies with serious methodological problems, necessarily weakening their conclusions.

Because of clinical implications, the relative and adjunctive benefits of medication and psychological therapies have long been debated by researchers. However, this debate is hampered by the limited evidence on which it rests: There is a relative paucity of large-scale trials, ands some earlier trials have been appropriately criticized for inadequate provision of medication, further reducing the pool of studies from which conclusions can safely be drawn. Reasonable expectations that trials ensure adequate dosage, and monitoring of drug compliance and prescriptions by adequately trained pharmacotherapists have not been met in many earlier trials, meaning that their inclusion in meta-analyses contributes confusion rather than clarity. Few trials contrast each treatment modality alone and in combination; given the number of possible combinations of psychotherapy and pharmacotherapy, it is inevitable that only some of the possible permutations have been explored. Furthermore, rather few trials include a contrast of active psychological therapy with pill placebo, a state of affairs that has led to some debate as to whether this does, or does not, weaken interpretation of current trials (e.g., Jacobson & Hollon, 1996a; Klein, 1990, 1996). All this hampers the inferences that meta-analysis can make about the relative and combined efficacy of these approaches.

Earlier reviews do not disaggregate particular types of psychotherapy, often contrasting any form of therapy against pharmacotherapy. The Quality Assurance Project (1983) largely focused on the efficacy of medication; of 200 trials in this review, only 10 involved the use of psychotherapy. On the

basis of this (limited) sample, an effect size of 0.69 was found for psychotherapy of any type, contrasted with a mean effect size of 0.55 for tricyclics and 0.39 for monoamine oxidase inhibitors (MAOIs). Conte et al. (1986) reviewed 17 studies published between 1974 and 1984, in which psychotherapy in combination with medication was contrasted with another treatment (all except two are included in Robinson et al.'s [1990] review). A statistical weighting procedure was used to reflect methodological adequacy, the results of which were used in the assessment of treatment effects. This procedure makes a number of assumptions about the link between methodological quality and outcome that have not been supported by more conventional meta-analyses (Smith et al., 1980), and its use may be questioned. No distinction among different types of therapy was made. Conte et al. (1986) conclude that there is some limited evidence for the greater effectiveness of combination treatments compared to psychotherapy or pharmacotherapy alone. Steinbruek et al. (1983) examined 56 studies published between 1962 and 1981, in which psychotherapy alone or pharmacotherapy alone was contrasted with a control group. A significantly larger mean effect size was found for psychotherapy (1.22) contrasted with medication (0.61). Although an effect size for tricyclic medications was derived (0.67), no separate analyses are given for the differing categories of psychotherapy. Furthermore, contrasting therapies against waiting lists would be expected to give a different result than comparison with placebo; in this study, both groups are subsumed under a single heading ("control group"). These considerations undermine the usefulness of this review.

Neitzel et al. (1987), and Dobson (1989) restricted their coverage to trials that used the BDI as an outcome measure, aiming to increase the uniformity of comparisons. Neitzel at al. (1987) analyzed data from 31 studies, concluding that the type of therapy, whether classified as cognitive, behavioral, or combination treatment, did not influence outcome and that individual therapy produced larger effect sizes than group treatment. To determine the clinical significance of treatment effects, they contrasted the mean pre- and posttreatment BDI scores of treated and control patients with the mean scores of two "reference groups" for which normative data for the BDI are available. The first set of norms is derived from individuals selected for the absence of pathology (a nondistressed group); the second set is norms derived from general population surveys. Contrasting outcomes against the former group is a more stringent (but perhaps less relevant) test of recovery than comparison with the latter sample. Using these contrasts, the average treated client moved from 4.79 standard deviations (*SD*) above the mean for the nondistressed group to 1.62 *SD* posttreatment; using general population norms, the comparable figures are 2.9 and 0.71 *SD*. Control subjects had similar pretreatment scores to those of treated patients but showed little gain posttreatment. (Their posttreatment scores were 3.97 and 2.96 *SD* when

contrasted with the nondistressed and general population samples.) Gains were maintained at follow-up, though it is worth noting that the mean follow-up period was only 16 weeks (range 4–52).

These figures can be expressed, perhaps more meaningfully, in terms of percentiles. On this basis, the average treated client moves from a pretreatment level at the 99.9th percentile to the 94.7th percentile of the nondistressed group at posttreatment. Contrasted to the general population, sample the gains are more striking—from the 99.8th percentile to the 76.1th percentile. Although there may be some debate about the clinical significance of these scores, they do indicate that the average treated client will have an approximate posttreatment BDI score of 10, which would place him/her at the boundary between a normal and a mildly depressed level of functioning.

Dobson's (1989) review was based on 28 studies published between 1976 and 1987, in which cognitive therapy was contrasted with other therapies, and (as with Neitzel et al., 1987) in which the BDI was included as an outcome measure. Given that Neitzel et al. did not cite the studies included in their review, the degree of overlap between the two analyses cannot be readily ascertained (though it is likely to be considerable). However, in contrast to Neitzel et al., Dobson (1989) found mean effect sizes favoring cognitive therapy contrasted with pharmacotherapy (0.53), behavior therapy (0.46), and other psychotherapies (0.54). Comparing cognitive therapy to wait-list or no-treatment controls gave a mean effect size of 2.15, indicating that the average treated client did better than 98% of control subjects. All effect sizes were based on posttherapy outcomes, with no analysis of follow-up data.

Robinson et al. (1990) reviewed studies published between 1976 and 1986, including 58 trials of outpatient psychological therapies and a further 15 in which psychological and pharmacological interventions were contrasted. Therapies were classified as behavioral, cognitive (in which treatments focused on the evaluation and modification of cognitive patterns), cognitive-behavioral, or general verbal. Only nine trials recruited patients through standard clinical channels; around half recruited using media announcements; in a quarter, the samples were students, and 12% did not report on the source of their patients. In addition, only 35% of investigations had inclusion criteria specifying that patients should meet formal diagnostic criteria for depression. The analysis is further weakened by the fact that most of the results are reported using posttreatment rather than follow-up data, on the basis that the figures for each data point were very similar (34 of the 58 studies had follow-up, but the mean length of follow-up was only 13 weeks, with a range of 2–52 weeks).

Contrasting therapy of any kind against no therapy (defined as wait list or placebo) gives an effect size of 0.73 at posttherapy and 0.68 at follow-up. As might be expected, effect sizes are higher when contrasted with a

wait list (0.84); when compared to placebo, the effect size is nonsignificant (0.28).

Overall, behavioral, cognitive, and cognitive-behavioral treatments showed moderate effect sizes (1.02, 0.96, and 0.85, respectively). Verbal therapies showed a more modest effect size (0.49), though this figure is based on only six studies.

Within-study treatment contrasts afford a more robust assessment of the relative efficacy of different classes of therapy. Using data from such trials, there is evidence of a modest superiority for cognitive-behavioral over behavioral treatment (effect size = 0.24), and a moderate superiority of cognitive and cognitive-behavioral treatments over verbal treatments (effect size = 0.47 and 0.27, respectively). Fifteen studies examined the implementation of these therapies in group or individual settings, finding equal efficacy for all approaches.

Longer treatment or the number of sessions did not contribute to a greater effect size, though the mean length of treatments was only seven sessions, and this may not be sufficient to test this variable. There was little evidence of any specific client characteristics that influenced treatment outcome. Effect sizes did not vary reliably when outcomes in relation to the initial severity of depression (as measured by BDI), the presence or absence of diagnosed depression, or the source of referral of the client were examined.

To examine the clinical significance of the gains made by patients (Jacobson et al., 1984), Robinson et al. (1990) identified 39 studies reporting data on the BDI, in order to contrast the level of depression after treatment with that obtaining in the general population (the same strategy followed by Neitzel et al. [1987]). On this basis, therapy appears to shift the average client from 2.4 *SD* above the mean for the general population to 0.8 *SD*. Comparison with a nondistressed population gives an equivalent pre- to posttherapy shift (from 3.4 to 1.4 *SD* above the mean) but does suggest some residual distress. Nonetheless, the clinical effect of treatment is clear.

Robinson et al. (1990) also examined 15 studies comparing the effectiveness of psychological interventions alone, medication alone, and the combination of these treatments. The majority of these studies employed cognitive-behavioral or behavioral psychotherapies. However, basic methodological requirements were not always met. Not all studies ensured that therapeutic levels of medication were prescribed, and although most trials utilized antidepressants, two (surprisingly) employed benzodiazepines. When contrasted with the use of tricyclic antidepressants, psychotherapy showed a significant though small advantage (effect size = 0.12). There was no evidence of a significant advantage to combination treatments contrasted with psychotherapy alone or medication alone. However, if the effect of investigator allegiance is controlled statistically (as discussed earlier), the advantage to psychotherapy disappears.

The AHCPR review (Depression Guideline Panel, 1993a, 1993b, 1993c) aimed to develop clinical practice guidelines for use by primary care practitioners and other health care professionals. However, most studies included in the analysis were conducted in tertiary care units, and it is unclear how research in these settings generalizes to treatment of the milder depressive conditions frequently seen in primary care. A later follow-up review by Schulberg et al. (1998, discussed below) considers this issue further.

The panel conducted an extensive review of studies published between 1975 and 1990, identifying a subsample for meta-analysis. Inclusion criteria specified that patients met DSM-III-R criteria for MDD, that the research design was a randomized controlled trial, and that treatment effects were measured using a standardized method that permitted assessment of depressive and/or functional status before and after treatment (Jarrett & Maguire, 1991; summarized in Jarrett & Rush, 1994). The number of included studies was further reduced by the decision to use Bayesian models rather than conventional meta-analytic methodology. Bayesian statistical methods can be described as a formal framework within which data from available studies are used to produce the best fit to a hypothetical, perfectly performed study. Existing data form a "prior distribution," which describes our present state of knowledge about, for example, outcomes for a particular therapy. The addition of further studies describes a "posterior distribution," representing our new state of knowledge as a consequence of the addition of further evidence. The application of Bayesian statistics results in a series of probability functions, which in essence describe the likelihood of a particular intervention having a particular result—for example, the probability that a treatment will be successful in resolving depression. It is argued that these methods are more sensitive and appropriate than conventional statistical techniques, and allow more scope for the management of possible bias in the studies forming the meta-analysis. Although the Confidence Profile Method (CPM) used in the AHCPR review (described fully in Eddy et al., 1990) can be undertaken using continuous data (e.g., differences in BDI scores), a categorical scoring system was used, because it has the advantage of indicating whether patients have reached a recovery criterion, hence making results more meaningful for practitioners.

Jarrett and Maguire (1991) identified 22 randomized controlled trials of cognitive therapy, 13 of behavioral therapy, 4 of interpersonal therapy, 8 of brief dynamic therapy, and 1 of marital therapy. Some of these studies were excluded because data were reported in a form that made it inimical to the CPM approach. Trials included in the analysis involved treatment of both adult and elderly populations, and used both individual and group formats for behavioral and cognitive therapy. Data were analyzed using "intention-to-treat" samples (contrasting outcomes for all patients entered into the trial, regardless of whether a full course of therapy was given).

All therapies combined yielded an estimated overall efficacy of 50%; contrasted to waiting list, there was a 26% advantage to therapy, and contrasted to placebo, a 16% advantage. The overall efficacy of behavioral therapy was 55%. Contrasted to all other forms of psychotherapy, it was 9% more effective (six studies), and compared to wait list (five studies), 17% more effective. Cognitive therapy was found to have an overall efficacy of 47%, with approximately equal efficacy to the other therapies (–4.5%, six studies). Brief dynamic therapy had an overall efficacy of 35%, though, contrasted to other therapies, it was slightly less effective (–8%, eight studies). For IPT, only the NIMH study provided data appropriate for meta-analysis; on this basis, it had an overall efficacy of 52% and exceeded the efficacy of cognitive therapy by 13%.

Comparison of all psychological therapies alone to medication showed an advantage to psychological therapy of 14% (based on eight studies). In all comparisons, specific therapies delivered alone showed advantage over medication alone: for behavioral therapy alone (24% more effective, based on two studies), cognitive therapy alone (15% more effective in three studies), brief dynamic therapy alone (8.5% more effective in two studies), and IPT alone (12%, based on the NIMH study). Combination treatment showed an advantage over behavior therapy alone (7% based on two studies) but was less effective than cognitive therapy alone (–6.5%, based on 6 studies).

The statistical sophistication of the CPM is both its strength and its weakness, because relatively few studies are appropriate for inclusion in this form of analysis. Studies meeting the inclusion criteria for conventional meta-analysis may not be representative of either clinical work or even of research trials, a criticism that may apply even more strongly to the CPM technique. This may lead to the emergence of misleading conclusions. Thus, few clinicians would accept that behavior therapy alone was the most effective treatment method for depression. Although achieving a high ranking in the meta-analysis of the depressions guidelines panel, this is most likely an artifact of sampling and of the six contrasts performed. Although CPM meta-analysis is an exciting development, it will require a considerable extension of the database to ensure that specific contrasts of therapies with one another can be interpreted reliably.

Gloaguen et al. (1998) located 78 controlled trials published between 1977 and 1996, in which cognitive therapy for depression or dysthymic disorder was contrasted to wait list, placebo, behavior therapy, or an alternative psychological treatment. Of these, 48 met inclusion criteria for methodological quality and were included in the review. Meta-analysis ostensibly demonstrated an advantage to cognitive therapy over wait list or placebo (20 comparisons, effect size = 0.82), over antidepressants (17 comparisons, effect size = 0.38), and over other therapies (22 comparisons, effect size = 0.24), and

equivalent impact against behavior therapy (13 comparisons, effect size = 0.05).

Though this appears to offer evidence of a strong advantage to CBT in contrast to other interventions, Gloaguen et al. (1998) noted significant heterogeneity across studies and further analyses. Wampold et al. (2002) suggest that this relates to a failure to separate contrast therapies into bona fide and non-bona fide therapies. This distinction relates to Wampold's (1997) observation that control treatments fall into two classes: bona fide treatments (which are expected to work) and non-bona fide interventions (whose impact is predicted to be negligible); clearly, contrasts between "active" treatments and non-bona fide interventions will overestimate relative treatment efficacy. Wampold et al. (2002) reanalyzed Gloaguen et al.'s (1998) data and found that separating bona fide from non-bona fide treatments restored homogeneity across studies. Crucially, when CBT was contrasted to bona fide interventions, outcomes were equivalent (effect size = 0.16).

The efficacy of group therapies has been evaluated by McDermut et al. (2001), who identified 48 trials published between 1970 and 1998. Most reports were of behavioral or cognitive-behavioral interventions, but eight were based on psychodynamic and/or interpersonal principles. Control conditions usually included wait list or delivery of the therapy on an individual basis. Contrast to no treatment (15 studies) yielded a mean posttreatment effect size of 1.03 in favor of group therapy. Comparing the same therapy delivered in individual and group formats (nine studies) yielded an effect size of –0.15, suggesting equivalence of outcome. Although dropout rates from individual and group therapies were very similar, there is evidence from two of the trials in the review that dissatisfaction with the group format was associated with greater attrition. This suggests that in routine clinical settings, it may be important to take into account patient preferences regarding the format of therapy delivery.

Westen and Morrison (2001) located 12 studies, published between 1990 and 1998, that met inclusion criteria for methodological quality, and included a comparison of an active psychological therapy with wait-list control, an alternative psychological therapy, or pharmacotherapy. Results are aggregated rather than identifying the specific benefits of specific approaches. Though this reflects Westen and Morrison's focus on broader issues relating to the presentation of outcome data, it limits the utility of the analysis for the present discussion. Though the mean pre- to posttherapy effect size for psychological therapies was large (effect size = 2.23) contrasted to control conditions, the median effect size at termination was small (effect size = 0.3).

Though 54% of patients who completed treatment met criteria for improvement, this dropped to 37% of the "intention-to-treat" sample; in line with other analyses, this raises issues about the proportion of patients who can

expect to benefit from therapy. Furthermore, in those studies that report relevant data, there may be cause for concern about the extent and stability of change. Although patients showed substantial gains, posttreatment mean scores on the HRSD and BDI remained above the conventional cutoff scores applied to these measures. These are usually set at ≥ 6 and ≥ 9, respectively; mean scores from Westen and Morrison's analysis (2001) were 8.68 and 10.98, respectively. Very few studies reported long-term follow-up data—only two at 12–18 months, and one at 2 years. At 18 months, only 36% of patients who completed treatment improved and remained improved; among the "intention-to-treat" sample, this figure drops to 28.5%.

Although only a preliminary finding, there was a suggestion that studies with higher exclusion rates prior to randomization to treatment tended to report better patient outcomes and lower rates of treatment seeking after therapy had ended. This raises a central dilemma for researchers. Setting clear inclusion criteria is a necessary and indeed helpful research strategy, but it has consequences for external validity. This is not a straightforward issue, however. Some exclusion criteria can be seen as enhancing external validity. For example, excluding patients who are only mildly depressed is both common and reasonable (though it also has the consequence of reducing information about treatment response in individuals whose depression lies below diagnostic thresholds). However, excluding patients with comorbid presentations, which is common in clinical settings, inevitably raises questions about generalization of results. Ultimately, this is an empirical rather than a philosophical issue, and one that will only be answered by more effectiveness trials, and particularly by more extensive use of benchmarking studies in routine clinical settings (e.g., Barkham et al., 2001).

Leichsenring (2001) reviewed evidence for the comparative efficacy of short-term psychodynamic psychotherapy (STPP) and CBT, restricting analysis to studies with more than 13 therapy sessions (a cutoff based on assumptions about the period required for psychodynamic therapy to show efficacy). Six studies were identified, though two of these employed therapies that may not best be described as either STPP or CBT. One of these was the NIMH trial, whose IPT arm is classified by Leichsenring as a psychodynamic intervention on the (accurate but contentious) basis that it was delivered by therapists whose background training was in psychodynamic therapy. In the second, Hersen et al. (1984) employed social skills training, a behavioral (rather than a cognitive-behavioral) treatment with unproven efficacy in relation to depression. Given these problems of definition, it is not quite clear that all trials contrast equivalent interventions. A final concern is that analysis of the Sheffield trials (Barkham et al., 1996a; Shapiro et al., 1994) is restricted to the 16-session arm, reflecting the dosage criteria described earlier. Broadly, in all but one trial, CBT and STPP (as defined by Leichsenring, 2001) were of equal efficacy. Gallagher-Thompson and Steffen (1994; reviewed elsewhere

in this chapter and in Chapter 15) looked at the impact of therapy on caregivers, finding some advantage to CBT for longer term (as opposed to shorter term) caregivers. Although the evidence from this meta-analysis is best taken to indicate an equivalence of action between CBT and STPP, this conclusion is based on only a small sample of direct contrasts of these approaches, itself enlarged by questionable inclusion criteria.

Churchill et al. (2001) reviewed brief psychological treatments for depression (defined as interventions of less than 20 sessions), contrasting each treatment with another and with TAU or wait list. Of 63 identified studies published between 1973 and 1998, 50 were suitable for meta-analysis. The authors noted that most samples were relatively small, with a range of 18–276 (median = 44), and a median arm size of 13. On this basis, few trials would have sufficient statistical power to detect differences between interventions. Only 60% of studies monitored adherence, and while 37% of trials requested that patients desist from medication, 20% allowed "naturalistic" prescribing—all trends that introduce ambiguity about the treatments received by patients. Only one-third of trials reported follow-up data, and those with follow-up of more than 6 months had high rates of attrition over this period. Although intended, a cost-effectiveness analysis was not undertaken, because only five trials reported relevant data.

Classes of therapy contrasted were variants of CBT (behavioral, cognitive, and cognitive-behavioral), psychodynamic therapy, IPT, and supportive therapy. The latter clustering included a mix of approaches, most of which were humanistic in orientation. Analysis was based on the proportion of patients recovered at posttherapy and (where available) follow-up. Compared to TAU or wait-list controls, there was significant benefit to psychological intervention. For all variants of therapy, the odds ratio (OR) was 3.01; significant benefit was apparent for CBT (OR = 3.42; 12 trials), IPT (OR = 3.52; one trial), and supportive therapy (OR = 2.71; 4 trials). Contrasts of CBT to commonly practiced alternative forms of therapy suggested equivalence with IPT (2 trials) but greater efficacy compared to psychodynamic therapy (OR = 2.23; 6 trials) and supportive therapy (OR = 3.45; 10 trials). CBT delivered in individual rather than group format was more effective in the short term (OR = 1.98; 6 trials), though this advantage was not evident at follow-up. No differences were found for the contrast of CBT and behavior therapy (3 trials). Based only on one trial, psychodynamic therapy showed equivalent efficacy to supportive therapy. Overall, while this analysis suggests consistent benefit to CBT, the number of available contrasts limit conclusions when contrasting the relative benefits of other specific therapies.

Several cautions are appropriate. Where follow-up data were available, between-therapy differences reported earlier were no longer evident. Studies in which patients were self-selected or volunteers tended to have larger effect sizes than those in which they were clinic attenders, and higher quality trials

also had lower effect sizes. It is also worth noting that there was significant heterogeneity in the data set contrasting CBT and other therapies; further analysis suggested that trials with more severely depressed patients showed fewer differences between therapies, and trials with fewer sessions favored CBT. Critically, there is overlap between this review and that of Gloaguen et al. (1998; reviewed earlier), and inclusion of contrasts of CBT and non-bona fide therapies may account for this heterogeneity, as well as inflating apparent between-therapy differences (Wampold et al., 2002). Finally, the reviewers inappropriately include the Sheffield trial in the contrast of CBT and IPT; although its presence or absence does not alter the significance of the odds ratio, its removal makes this comparison reliant solely on the NIMH trial.

"Mega-Analyses"

Thase et al. (1997) presented a "mega-analysis" of six studies contrasting the impact of psychological therapy alone or in combination with medication. All trials in the analysis were conducted completely or partially at the University of Pittsburgh (hence, including the NIMH and Frank studies reviewed earlier). In total, 243 patients received psychological therapy alone (CBT or IPT), and 352 received IPT in combination with antidepressant (imipramine or nortriptyline). Less depressed individuals (those having an initial HRSD score of ≤ 19) had equivalent outcomes with either treatment modality, but for more severely depressed patients, there was a significant advantage to combination treatment over psychological therapy alone (with respective recovery rates of 43% and 25%). For patients with recurrent depression, this advantage was even more marked (60% vs. 19%). A sensitivity analysis indicated that this pattern of results was consistent across all studies, further supporting the contention that patients with severe and recurrent depression will benefit from combination therapy.

DeRubeis et al. (1999) examined a different set of contrasts—the relative benefits of CBT alone and pharmacotherapy alone (imipramine or nortriptyline). The analysis combined data from four trials—the NIMH study (also included in Thase et al., 1997), the University of Minnesota trial, Rush et al. (1977) and Murphy et al. (1984). Data were reanalyzed to include only the more severely depressed patients (those with an initial score > 20 on the HSRD or > 30 on the BDI). Comparing CBT and medication using the HRSD yielded a small effect size of 0.22 in favor of CBT; when the BDI was used, there was no difference in efficacy between treatment modalities (effect size = 0.07). Unlike the Thase analysis, there is some variation in the pattern of results across studies, with better outcomes for medication and poorer outcomes for CBT in the NIMH trial when contrasted to the pooled outcomes from the other trials (though the effect size of approximately 0.5 was not significant). In contrast, the Rush et al. (1977) study showed an inverse pattern,

with outcomes for CBT being markedly better than those for medication (though, as detailed below, this study has been criticized for poor medication delivery).

Casacalenda et al. (2002) identified six studies that contained direct contrasts of medication, psychotherapy, and a control condition, and that contained data on remission (defined as a score on a depression scale within the normal range) (Elkin et al., 1989; Herceg-Baron et al., 1979; Jarrett et al., 1999; Mynors-Wallis et al., 1995; Schulberg et al., 1996; Scott, 1992). Medications in these trials were either tricyclics or phenelzine, and psychotherapies were usually CBT or IPT; most patients had mild to moderate levels of depression. Remission rates with active treatment were equivalent at approximately 46% of the intent-to-treat sample in contrast to 24% for control conditions.

A critical issue in these analyses is the relative impact of treatments for more severely depressed patients. In this respect, it seems reasonable to conclude that while the evidence for differential efficacy of each modality delivered alone is not strong and appears mixed, the evidence for the benefit of combination treatment over therapy alone appears both stronger and more consistent.

Researcher Allegiance and Outcome

Robinson et al. (1990) employed a 5-point scale to estimate the degree to which researchers displayed an allegiance to the therapy under examination and found a significant correlation between this measure and the results of direct comparisons between treatments ($r = .58$). Using a regression analysis to partial out this effect reduced the effect sizes almost to zero. Gaffan et al. (1995) explored this important issue further by reanalyzing the same sample of studies examined by Dobson (1989), and extending their sampling by including a further 37 studies published between 1987 and 1994. Their analysis of Dobson's original sample yielded similar effect sizes and conclusions, though controlling for sample size reduced the magnitude of the effect sizes. The allegiances of researchers were strongly biased in favor of cognitive therapy, and regression analysis suggested a significant association between this measure and estimations of outcome. Although controlling for this reduced the effect size markedly, the superiority of cognitive therapy over other treatments claimed by Dobson remained (corrected effect size = 0.17). Analysis of the second (more recently published) sample yields a set of conclusions similar to Dobson's sample. Though effect sizes are smaller, sample sizes are larger, and weighted and unweighted effect sizes are not significantly different. Strikingly, regression analysis suggested no association of investigator allegiance and reported outcome. Considering the complete sample of trials, year of publication emerges as a significant factor in mediating the link

between effect size and allegiance, with early trials (most obviously those published up to 1985) showing the strongest influence of allegiance. Understanding this shift is not easy: It would seem that both the effect size favoring cognitive therapy and allegiance have declined with time.

Summary

Summarizing quantitative reviews of outcomes for MDD, it seems clear that psychological therapy has benefit over no therapy, though when active therapies are contrasted, differences between them are less clear. Although there are indications that CBT is superior to less structured forms of psychotherapeutic intervention, it is worth noting that this conclusion appears less robust when the contrast treatment is credible and theory-grounded. There are few well-structured studies of psychodynamic therapy, though in the one, perhaps limited review in which this approach was explicitly contrasted to CBT, reasonably similar outcomes were evident. Group therapies appear to have comparable effect sizes when individual and group-based forms of the same approach are compared. Though persons with depression are vulnerable to relapse, relatively few studies include adequate follow-up data, making it difficult for reviews to derive conclusions about the robustness of change. When they do, they confirm the lack of stability of posttreatment gains.

Estimates of effect sizes differ across meta-analyses as a function of the sampling frame of the original studies, the dependent measures used, the statistical principles of the meta-analysis, and the original date of the investigation. Broadly speaking, larger effect sizes are observed when trials include less severely impaired and more highly selected patient samples assessed on clinician ratings as opposed to self-report measures, when they employ shorter rather than longer follow-up, and (though perhaps less obviously in more recent studies) when they are carried out by the proponents of techniques. A critical methodological concern should be to distinguish between bona fide or non-bona fide control treatments, since contrast of an active therapy to plausible and theoretically grounded techniques appears to reduce effect sizes.

Though a critical clinical issue, judging the relative benefits of medication and psychological therapy offered either alone or adjunctively is difficult. There is no consistent evidence that psychotherapeutic treatment is more effective than pharmacological treatment, but there is less certainty as to whether medication has benefit over psychological therapy, and especially whether there is an advantage to combination treatment. In general, there has to be concern that at least some studies included in reviews are fundamentally flawed in their implementation of medications. It is also the case that few of the trials available to reviewers employed newer, more specific antidepressants (with fewer side effects and probable greater acceptability to patients) that are likely to have lower rates of attrition.

While the strength of meta-analytic reviews is their capacity to identify trends that individual studies do not, aggregation of data often relies more on methodological than clinical considerations. This can result in contrasts that disguise important clinical differences between studies, obscuring clinically pertinent issues and yielding erroneous or misleading conclusions concerning treatment effectiveness. Examples of this would be the comparison of active treatments to ones that are not expected to be effective (hence boosting apparent efficacy), or when effect sizes for an intervention are based on trials where it was implemented as a control treatment rather than an "active" therapy (in this way probably reducing its impact).

SPECIFIC TREATMENT APPROACHES

Pharmacotherapy and Psychotherapy in Combination and Alone

This section supplements the quantitative reviews already discussed with qualitative reviews, and considers individual studies because they have either been published since these reviews or they illuminate the current status of an approach.

A number of methodological concerns complicate interpretation of trials that contrast these two modalities.

1. *Inadequate implementation of medication*. Though recent trials are more attentive to the need to implement pharmacotherapy in a manner that its proponents would recognize as appropriate, earlier researchers sometimes failed to ensure that patients received adequate dosages, did not conduct checks on compliance, and in some cases used inappropriate medications (such as anxiolytics), creating non-bona fide contrasts and almost certainly enhancing psychotherapy treatment effects. Meterissian and Bradwejn (1989) examined 11 studies carried out between 1977 and 1987, finding that only around half employed optimal levels of antidepressants, and only three measured drug plasma levels. Two widely cited studies illustrate the problems of inadequate prescription of medication. Rush et al. (1977) employed low doses of medication, which were then rapidly withdrawn before outcome measures were taken. Blackburn et al.'s (1986) study was carried out in both an outpatient and a primary care setting; patients in primary care showed an unusually low response to medication, against which therapy inevitably showed a superior outcome.

2. *Inadequate representation of newer antidepressants*. Many studies of medication utilize "older" medications—usually tricyclics—rather than SSRIs. The import of this is unclear. Although SSRIs have relatively fewer side effects and are hence likely to be better tolerated by patients, evidence that they are more effective than older medications is equivocal (e.g., Barbui & Hotopf, 2001).

3. *Influence on outcome of "blindness" to medication status.* Detecting the impact of medication requires control for patient and researcher expectancies, since these are known to influence outcomes. Although most medication trials attempt to ensure that all parties are "blind" as to the medication received, this is hard to achieve. Patients receiving placebo are less likely to experience side effects or other indicators that they received active medication, and this will impact on their expectancies of outcome. Whether consciously or not, clinicians making judgments about outcome will also use the presence of side effects as a way to determine treatment allocation. There is also good evidence that patients, as well as therapists, are able to detect whether they were allocated to drug or placebo, despite attempts at "blinding" (e.g., Margraf et al., 1991). More subtle effects are also apparent. Greenberg et al. (1992) identified 22 trials in which both new antidepressants were contrasted to placebo, with older antidepressants also used as a control treatment, arguing that researchers would have less investment in the older antidepressants and (because two active treatments were under test) would find it harder to distinguish between the two. Under these conditions, there was a marked reduction in effect size relative to control: Effect sizes for medication fell to between one-half and one-fourth of those usually reported, suggesting that—in some way—expectations of their efficacy influenced apparent outcome. Further evidence for expectancy effects comes from studies employing "active" placebos (which mimic drug side effects). Of nine such trials, only two showed effect sizes significantly favoring active medication (Moncrieff et al., 1998). These reports raise questions about the potential for overestimating drug effects and—perhaps more pertinent to this review—the possibility of bias when contrast of medication and psychotherapy is undertaken by researchers with an allegiance to either modality. These points also reinforce Hollon and DeRubeis's (1981) argument that psychotherapy alone is not equivalent to psychotherapy plus placebo, largely because of the expectancy effects associated with the latter combination. On this basis, they suggested a nine-cell design for medication and psychotherapy comparisons in which all possible combinations of drug, placebo, and nonintervention could be tested. While this represents an ideal rather than a practical suggestion, it does point to the difficulty of interpreting the more usual restricted set of contrasts.

4. *Depression subtype and response to placebo control.* There is an ongoing (and at points contentious) debate regarding the classification of depression, with the suggestion that the "endogenous" subtype reflects a biological etiology responsive to medication but not to psychological intervention (e.g., Feinberg, 1992; Thase & Friedman, 1999). It is also argued that patients with milder and, hence (at least potentially), more transient cases of depression show a greater response to most nonspecific interventions. On this basis, no active therapy would be expected to show benefit over another, leading to an erroneous conclusion that all therapies have equal efficacy.

Reflecting these concerns, some researchers (e.g., Klein, 1996) have argued that studies contrasting psychotherapy and medication should include a placebo control to demonstrate that the sample was (as it were) medication-responsive—by implication more severely distressed, with a low rate of placebo response, and therefore likely to be treatable only with the application of specifically effective therapies. However, there is a risk that these arguments become self-fulfilling, drawn on selectively to critique trials in relation to outcomes rather than criteria for trial entry. Though, on this basis, the specific argument may be overstated, there is merit in this position, though not simply in relation to contrast of medication and psychotherapy. It seems reasonable to question the degree to which maintaining diagnostic homogeneity protects researchers against predictable variations in treatment responsiveness within any sample of depressed individuals. However, the challenge is to derive a reliable and valid system for factoring any such differences into clinical trials.

5. *Controlling for the impact of attention*. When contrasted with patients receiving medication alone, psychotherapy patients are likely to receive three or four times more time with their therapists (typically, three or four times as much time), making it hard to distinguish the relative benefits of attention as opposed to therapeutic interventions. Again, this issue is managed better in at least some recent studies, where regular medication clinics ensure regular (if nonspecific) therapeutic contact.

Three qualitative reviews cover the period up to the early 1990s, inevitably including rather few contrasts of psychotherapy and newer medications. Meterissian and Bradwejn's (1989) review (described in part earlier) identified 11 studies, of which five were considered to have offered adequate doses of medication. Of these, one reported psychotherapy to be superior to medication; the remainder found it equivalent. In those studies using inappropriately low doses of medication, two studies indicated that psychotherapy was superior, one that it was equivalent, and one that a combination treatment was best.

Wexler and Cicchetti (1992) examined outcomes from eight treatment studies; Manning et al.'s (1992) review is rather larger, including 17 trials (six of which were included in Wexler and Cicchetti).

All studies in Wexler and Cicchetti's (1992) review used the BDI as a measure of outcome and employed broadly similar criteria for recovery. Psychological treatments included behavior therapy, CBT, psychodynamic therapy and IPT; medication was invariably a tricyclic. Three trials compared psychotherapy alone, pharmacotherapy alone, and the two in combination. Two of these (Blackburn et al., 1981; Murphy et al., 1984) suggested that combination treatments had the greatest efficacy, with intermediate efficacy for psychotherapy, and the lowest for medication. The third (Hersen et al., 1984) showed a nonsignificant trend favoring psychotherapy over combina-

tion, with medication showing the least efficacy. Combined data from these studies suggested a trend indicating that combination treatments were the most effective, followed by psychotherapy alone and medication alone. Overall, however, there was no statistically significant evidence demonstrating the greater efficacy of any one treatment over the other two.

In three studies comparing psychotherapy alone to medication alone (Elkin et al., 1989; McLean & Hakstian, 1979, 1990; Rush et al., 1977). two showed an advantage to psychotherapy, while in the NIMH trials, similar response rates were obtained. Finally, Beck et al. (1985) contrasted psychotherapy alone to combination treatment, producing results favoring psychotherapy over medication. However, these studies may form a poor basis for comparison of medication and psychotherapy. The Rush et al. (1977) study has been criticized for poor implementation of medication regimens; Beck et al. (1985) employed relatively low doses of amitriptyline for short periods; and in the NIMH (Elkin, 1994) trial, medication was offered in combination with clinical management.

Overall, Wexler and Cicchetti (1992) estimated that of 100 patients treated, the success rate will be 29% for medication contrasted with 47% for both psychotherapy alone and combination treatments. These figures suggest a strong advantage for psychotherapy or its combination with medication. However, it is clear that each therapy modality shows a high rate of partial or nonresponse—42% with psychotherapy alone and 52% with medication alone, suggesting that treatment strategies based exclusively on either modality will show little benefit to a significant number of patients.

Manning et al.'s (1992) review included contrast of psychoanalytic, behavioral, cognitive, interpersonal, and marital therapies, usually contrasted with tricyclic medication, though in two cases with benzodiazepines. The reviewers note that no studies were conducted by investigators with a primary allegiance to pharmacotherapy. Small sample sizes (hence, low statistical power) were identified as a problem by these reviewers: Only two studies had more than 30 patients per treatment cell. In addition, all but two trials focused on the acute phase of antidepressant treatment. Only seven included follow-up data (varying from 3 to 24 months), and four focused on maintenance therapy.

The efficacy of therapies was examined using a "box score" method contrasting the relative outcomes associated with each treatment modality rather than their absolute impacts. Although some advantage was found to combination treatment, no clear superiority was evident; however, in no study was combined therapy less effective than its component treatments. Summing across studies, combined treatments outperformed medication alone in 4 of 10 instances (40%) and outperformed psychotherapy alone or with placebo in 7 of 18 instances (39%). Not enough trials were available to examine the impact of specific psychotherapies.

Data from a number of studies have suggested that both treatment failure and rate of dropout were highest for patients receiving medication. Given that the cited success rate of most studies usually reflects treatment completers, these factors are important to any consideration of clinical effectiveness. There is some evidence to suggest that dropout reflects dissatisfaction with treatment rather than clinical improvement. Weissman et al. (1979) reported that 92% of patients drop out because of dissatisfaction with treatment; in the NIMH study (Elkin et al., 1989) 77% of dropouts did so for similar reasons. Fawcett et al. (1989) reported that 98% of dropout was because of problems with medication side effects or dissatisfaction with treatment. Wexler and Cicchetti (1992) estimate the rate of dropout at 52% for medication, 30% for therapy alone, and 34% for combination treatments. Casacalenda et al. (2002) reported attrition at 37% and 22%, respectively, for medication and psychotherapy. In contrast, Manning et al. (1992) reported that of 11 studies reporting attrition rates, only one study found that attrition rates varied between single and combined treatments.

Psychodynamic Therapy and Medication

Two recent studies examined the impact of adding psychodynamic therapies to medication, though both are somewhat problematic. de Jonghe et al. (2001) contrasted pharmacotherapy alone or in combination with "short-term psychodynamic supportive therapy." Psychotherapy was delivered in 16 sessions over 6 months, but although described as psychodynamically informed, the protocol emphasized supportive and problem-solving elements, and excluded the use of transference interpretations (raising questions about the best description for the therapy being conducted). Pharmacotherapy employed fluoxetine in the first instance; dependent on patients' responses, the protocol allowed for them to be switched to amitriptyline or (finally) to moclobemide. Of the 167 patients randomized to treatment, high initial refusal rates meant that 57 started pharmacotherapy and 72 started combined therapy. At the end of the 6-month treatment period, and with remission defined as an HRSD score ≤ 7), combined treatment showed significant advantage over pharmacotherapy alone (21% and 44% of the completer sample achieved remission at posttherapy), with evidence that more rapid remission was achieved with the combined treatment. However, at later stages of the trial, there was significantly greater attrition from pharmacotherapy alone contrasted to combined treatment, which may have contributed to this finding.

Burnand et al. (2002) assigned 95 patients with HRSD scores ≥ 20 to 10 weeks of either psychodynamic therapy in combination with clomipramine, or the same period of treatment with clomipramine alone. In practice, clomipramine was combined with "supportive care" offered at the

same intensity as psychodynamic therapy, and comprising "empathic listening, guidance support, and facilitation of an alliance." Psychotherapy was offered by nurses working under the supervision of a psychoanalyst. Both groups evidenced equivalent gains at 10 weeks, though the proportion meeting criteria for MDD was lower among those receiving psychodynamic therapy. In contrast to the rate of response found in most trials, improvements were unusually marked and rapid for such a brief period of intervention: HRSD means in both treatment groups reduced from a mean of approximately 24 to 9 at posttherapy, perhaps raising questions about the sample under study. After this acute treatment phase, patients were rereferred to routine outpatient treatments, though at clinicians' discretion, some continued in the original treatment condition to which they had been randomized (Burnand, personal communication, 2003). Monitoring continued until clinicians discharged patients (based on clinical rather than research criteria), though variations in the length and type of additional treatment make it difficult to attribute longer term outcomes to the original treatment received. With this important caveat, there was some evidence that patients who received psychodynamic therapy had fewer inpatient admissions, suggesting some modest advantage of psychodynamic therapy in combination with medication when contrasted to medication combined with counseling and support.

Cognitive-Behavioral Therapy and Medication

In a trial that also examined the impact of maintenance therapy, Blackburn and Moore (1997) allocated 75 patients to one of three combinations of acute–continuation treatments—pharmacootherapy in both phases, CBT in both phases, or acute-phase pharmacotherapy followed by CBT. Acute treatment lasted for 16 weeks and the continuation phase, for 2 years. Patients were relatively unselected contrasted to many trials, and many had both chronic and severe presentations. Medication was prescribed under naturalistic conditions through hospital or primary care doctors (hence, patients received a range of antidepressants), though a research protocol specified minimum dosages. Patients receiving CBT met with their therapists weekly in the acute phase and monthly thereafter; no comparable information on frequency of contacts is available for the medication group, though, on the basis of a naturalistic mode of delivery, it probably varied widely. All three groups showed equivalent gains across both phases of the trial, though (perhaps reflecting the sample) overall response rates were somewhat lower than for comparable trials (using a criterion of an HRSD score ≤ 6; rates of remission after the acute phase for (pooled) pharmacotherapy and CBT were 24% and 33%, respectively).

Jarrett et al. (1999) contrasted the efficacy of cognitive therapy, phenelzine, and placebo in 108 patients with "atypical" depression (a subclass of DSM-IV-TR MDD usually associated with a chronic and recurrent course). The choice of an MAOI was based on evidence of greater efficacy of this medication with this population. Over 10 weeks of treatment, cognitive therapy and phenelzine appeared to have equivalent efficacy, and both were more effective than placebo. A recovery criterion of an HRSD score ≤ 9 was met by 58% of patients in both the cognitive therapy and phenelzine groups contrasted to 28% of those receiving placebo.

A large multicenter trial conducted by Keller et al. (2000) randomly assigned 681 patients to one of three conditions—nefazodone alone, a variant of CBT, or a combination of the two treatments. Although the psychotherapy employed—cognitive behavioral analysis system of psychotherapy (BCASP)—contains many standard elements of CBT, it has a strong interpersonal focus, emphasizing social problem solving and making use of interpersonal themes arising in the therapeutic relationship. Patients were recruited on the basis that they had a chronic condition, with their current episode of MDD lasting for at least 2 years. Psychotherapy was delivered twice weekly for the first 4 weeks, and weekly thereafter to 12 weeks. Remission was defined as an HRSD score ≤ 8 at weeks 10 and 12. On this basis, psychotherapy alone and medication alone had equivalent outcomes (33% and 29%, respectively), but combination therapy showed significant advantage both in terms of remission (48%) and in the speed of recovery.

Interpersonal Psychotherapy and Medication

Reynolds et al. (1999a) reported on outcomes in 187 patients over 60 years of age, recruited over a 7-year period on the basis that they presented with a current episode of MDD, had a HRSD score ≥ 17, and had experienced at least one prior episode of MDD. In the first acute phase of the trial, all patients received a combination of nortriptyline and weekly sessions of IPT, until they achieved remission (defined as an HRSD score ≤ 10). Following remission, patients received 16 weeks of continuation treatment; 107 patients met criteria for a stable recovery and entered a maintenance phase (which lasted for 3 years, or until they relapsed). At this point, they were randomized to receive one of four possible treatments: medication clinic and nortriptyline alone; medication clinic and placebo alone; monthly maintenance IPT with nortriptyline, or monthly maintenance with IPT alone. Survival analysis over 3 years indicated a clear benefit to all active treatments over placebo, and for the combination of nortriptyline and IPT over other active treatment options. Recurrence rates for the combination of nortriptyline and IPT were 20%; for nortriptyline alone, 43%, for IPT with placebo, 64%, and for placebo, 90%.

A second report by the same research group (Reynolds et al., 1999b) focused on the efficacy of nortriptyline or IPT on later life depression associated with bereavement. Eighty patients over 50 years of age were recruited on the basis that their depression began close to the death of a partner (either 6 months prior to the death or in the year following it), and randomly assigned to one of four treatments: IPT combined with nortriptyline, IPT combined with placebo, nortriptyline alone or placebo alone. Reflecting trial entry requirements, patients were largely self-referrals and most were female. The protocol allowed for 8 weeks of acute treatment under double-blind conditions; at this point, patients in remission (defined as having an HRSD score ≤ 7 over 3 consecutive weeks) were entered into a 16-week continuation phase. The trial protocol allowed clinicians to break the blinding conditions if they were concerned about patients' progress during the acute phase, effectively compromising the initial randomization; there was a marked difference in the mean length of the acute phase in each "arm" of the trial. Patients receiving nortriptyline and IPT completed a mean of 76 days of treatment contrasted to (approximately) 50 days in each of the other contrasts (patients in IPT combined with nortriptyline received a mean of 9 sessions and 6 when IPT was combined with placebo). Remission rates for the combination of nortriptyline and IPT were 69%; for nortriptyline alone, 56%; for placebo, 45%; and for placebo plus IPT, 29%. At face value, this provides robust evidence for the benefits of active medication and at best suggests some benefit to its combination with IPT. However, some caution seems appropriate because of the methodological problems discussed earlier: the small sample sizes in each treatment cell and the unusually high placebo response rate.

Frank et al. (2000) described the use of two different strategies for implementing IPT and medication in a successive cohort design that recruited women in their second or greater episode of MDD. In the first (sequential treatment) condition, 158 women received IPT alone for between 12 and 24 weeks until remission, followed by a further 17 weeks of IPT. If continued remission was not evident at this point, IPT was combined with an SSRI. In the second (combination) condition, IPT and medication were offered together from the outset, with a broadly similar pattern of psychotherapy. Remission rates in 159 women offered sequential treatment were 50% with IPT alone, boosted to 79% with the addition of medication. In 180 women offered combination treatment, remission rates were 66%. Lack of randomization and small differences in delivery of IPT caution against overinterpretation, particularly because a theoretical rationale for the observed differences in remission rate is not obvious. Nonetheless, this does suggest that the manner in which combination treatment is initiated may be pertinent to outcome.

Couple Therapy and Medication

Based on evidence that the degree of criticism expressed by a partner toward a patient is associated with a poorer prognosis, Leff et al. (2000) recruited depressed patients living in a stable relationship with a critical partner. Seventy-seven patients were stratified into those with and without a previous history of significant depression, and were then randomized either to receive couple therapy or pharmacotherapy. Though desipramine was prescribed initially, nonresponders were prescribed trazodone or fluvoxamine; blood testing was used to check adherence. Alongside medication, psychoeducation and 12–20 outpatient sessions were offered over 1 year, after which medication was tapered. Couple therapy was conducted using a flexible protocol that identified and attempted to remedy problematic patterns of interaction, and was delivered by experienced clinicians over 12–20 sessions. At 1 and at 2 years, intention-to-treat analysis showed gains for patients in both groups, though whether mode of therapy impacted differentially was dependent on the form of measurement: The BDI indicated a significant advantage to couple therapy, but on the HRSD, both therapies showed equivalent outcomes. This rather ambiguous outcome is difficult to interpret, but a conservative interpretation suggests that couple therapy is of equivalent efficacy to medication for individuals with histories suggesting that relationship issues may be relevant to their presentation.

"Problem-Solving" Treatment and Medication

Problem-solving treatment has a psychosocial, here-and-now focus, and encourages patients to specify and work toward resolving areas of functioning that they identify as problematic. Most trials examine the impact of this approach for individuals with mild to moderate depression in primary care settings.

Mynors-Wallis et al. (1995, 2000) recruited patients through primary care centers in Oxfordshire. They (1995) allocated 91 patients to one of three treatments—amitripyline and clinical management, placebo and clinical management, or six sessions of problem solving delivered over 12 weeks. Setting a recovery criterion as HRSD scores ≤ 7 at 12 weeks, the recovery rate in patients receiving amitripyline and problem-solving therapy was equivalent (60% and 52%, respectively), but was significantly greater than for patients receiving placebo (27%). A later study by the same research group (2000) randomized 151 patients to receive problem solving therapy alone, fluvoxamine or paroxetine alone, or problem solving combined with one or the other of these medications. Over 12 weeks, all treatments showed equivalent efficacy. It should be noted that entry criteria for these studies are lower than

in most trials reported in this section (patients needed to meet research diagnostic criteria for depression and have an HRSD score ≥ 13, usually an indicator for mild depression). On this basis, it may be inappropriate to generalize these results beyond benefit to patients with mild to moderate depression.

Brief Dynamic Therapy

There continue to be fewer controlled trials of brief dynamic therapy than would be expected given its widespread use in clinical practice. Unfortunately, of those studies available, few have been carried out by proponents of the technique, and often dynamic therapy has been employed as a contrast to alternative therapies with which the investigators were professionally identified. Treatment periods are usually short [a mode of 12 sessions (range = 12–36) in the studies reviewed in this chapter], which may be too short for this technique. Therapists in these trials were unlikely to administer dynamic therapy appropriately because of their lack of commitment to the method. These methodological problems, together with the likely bias introduced by investigator allegiance, suggest that results from these studies should be viewed with caution.

Many studies of psychodynamic therapy suggest that dynamic therapy is significantly less effective than other forms of intervention. Thus, Steuer et al. (1984) found it less effective than cognitive therapy; McLean and Hakstian (1979) found that it performed more poorly than behavior therapy; and Covi and Lipman (1987) found that it was less effective than both cognitive therapy alone and cognitive therapy combined with medication. Kornblith et al. (1983) contrasted behavioral self-control methods against dynamic therapy; all treatments were administered in groups and were found to be equally effective. However, small sample sizes and variations in sample size across treatment conditions make interpretation of this study difficult. Bellack and colleagues (1981; Hersen et al., 1984) treated 50 depressed women with amitriptyline, social skills training and medication, social skills training and placebo, or dynamic therapy with placebo (designated as a "nonspecific therapy"). All treatments resulted in equivalent gains. Thompson et al. (1987, 1990) contrasted dynamic therapy, cognitive therapy, and behavioral therapy against a wait-list control; all treatments were delivered in group formats and with elderly depressed patients. Dynamic therapy was more effective than a wait-list control, and all three treatments were equally effective both posttherapy and at 1- and 2-year follow-up.

Although outcomes from the Sheffield psychotherapy study (Shapiro et al., 1994) are a more robust demonstration of the potential efficacy of this technique, it remains the case that, overall, support for brief dynamic therapy is sparse and at best equivocal. No study favors dynamic therapy over other therapies, and some suggest that it performs more poorly. Firm conclusions

regarding the efficacy of brief dynamic techniques still require further and better designed research.

Interpersonal Psychotherapy

An unusual—indeed, probably unique—trial reports on the application of IPT in rural Uganda (Bolton et al., 2003). Using a cluster randomized design, 30 villages were selected for study. Because therapy was conducted in gender-specific groups, in 15 of these villages participants were male, and in the remaining 15 they were female. Half the "male" and "female" villages were assigned to the intervention, with half acting as a control. Potential participants were identified by local leaders and screened using appropriately adapted standardized measures. Trial entrants were also required to meet DSM criteria for MDD, though these were slightly relaxed, allowing entry for individuals who fell short of MDD diagnosis by one symptom criterion. In each village, IPT was offered in a group format over 16 weeks, with 116 villagers receiving this intervention and 132 acting as controls. Therapists were locally recruited and trained, and appear not to have had prior psychotherapeutic expertise. Contrasted to controls, IPT resulted in a significant reduction in symptoms and a marked reduction in rates of diagnosed MDD (which fell from 86% to 6.5% in the IPT group, and from 94% to 55% in controls). This carefully constructed study demonstrates the benefits of a psychosocial intervention (though not necessarily the specific benefits of IPT) despite clear cultural differences between villagers' conceptualizations of depression and those of the patients for whom IPT was originally developed.

Couple Therapy

Baucom et al. (1998) and Beach et al. (1998) reviewed four studies that contrasted individual IPT and couple-based IPT (Foley et al., 1989), behavioral marital therapy, individual cognitive therapy or a combination of both approaches (Jacobson et al., 1991), behavioral marital therapy, individual cognitive therapy or a wait-list control (Beach & O'Leary, 1992), and individual cognitive therapy and "communication-focused marital therapy" (Emanuels-Zuurveen & Emmelkamp, 1996). These trials were consistent in indicating that, for depressive symptoms, there was an equivalence of action between individual and couple therapies, and between variants of couple therapy. However, when the focus of the therapy lay more with discordant marital relationships, there was evidence that couple therapies were more effective than individual therapy. A caution on the generalizability of these results is in order, because all trials had relatively small sample sizes, and the majority of index patients were female. Leff and Everitt (2001, reviewed earlier) found that pharmacotherapy and couple therapy were of equivalent efficacy for

couples selected for marital distress. Current evidence does not provide robust indicators for couple therapy as contrasted to individual therapy. Nonetheless, there is some support for the clinically intuitive notion that couple-based approaches are preferable when relationship stress is a prominent feature of the presentation, because of their differential impact both on the quality of the relationship and on symptoms.

Bibliotherapy and Computer-Aided ("Self-Help") Therapy

These interventions involve little or no direct therapist–client contact, and the usual contrasts are to the efficacy of a similar intervention offered by a therapist. Only behavioral and CBTs have been adapted to this form of delivery; these modalities have a clear rationale that can be presented in a systematic and structured manner. Nonetheless there are significant variations in the way the "self-help" version of the therapies is delivered; at one end of the scale, studies examine the impact of reading recommended texts; at the other, patients interact with a sophisticated, computer-aided therapy package that tailors itself to individuals' needs. Because there is relatively little research in this area, the heterogeneity of approaches suggests that conclusions about the efficacy of self-help approaches need to pay due regard to the specific interventions employed. Cuijpers (1997) conducted a meta-analysis, and Williams and Whitfield (2001), a qualitative review, of bibliotherapy and computer-based treatments.

Bibliotherapy

Cuijpers' (1997) meta-analysis identified seven trials of bibliotherapy for depression, of which six met criteria for methodological quality. Bibliotherapy materials employ a behavioral or cognitive-behavioral approach, and all included at least some therapist contact, though the extent of this varied widely from trial to trial (three had weekly contact, one had contact at the beginning and end of the "session," one at the start and finish of the trial, and one at the start, middle, and end point). Against wait-list control, there was a mean effect size of 0.82; against individual therapy (four studies) the effect size was 0.1. Although this suggests that there may be utility to this approach, all patients were nonclinical populations recruited through media announcements.

Scogin et al. (1989) used media announcements to recruit 67 mildly to moderately depressed older adults randomized to receive one of two forms of bibliotherapy or to be placed on a wait list. The two bibliotherapy texts set out either a behavioral or a cognitive-behavioral model of depression management; participants read the books over a period of 4 weeks, with a research assistant phoning at weekly intervals. Both forms of bibliotherapy

resulted in a significant reduction in BDI scores contrasted to the wait list; after treatment, the wait-list group showed improvements equivalent to those receiving immediate treatment. Two-year follow-up (Scogin et al., 1990) suggested that gains were maintained.

Using a similar research design but with a younger sample, Jamison and Scogin (1995) randomized 80 patients solicited through newspaper advertisements to bibliotherapy or to a wait-list control. Bibliotherapy employed a book setting out a CBT model of depression; again, a research assistant phoned at weekly intervals over the 4-week intervention period. Posttherapy patients receiving bibliotherapy showed a significant reduction in BDI scores; at this point, only 30% of patients met DSM criteria for depression, contrasted to 97% of wait-list controls. After the wait-list group received bibliotherapy, their outcomes were equivalent to the immediate treatment group. Three-year follow-up (Smith et al., 1997) found that gains were maintained, although 44% of those followed up had either sought further help for depression or met criteria for depression at the time of follow-up. This result, though impressive, exceeds the usual epidemiological pattern of relapse in depression; hence, it raises some question about the representativeness of the sample.

Beutler et al. (1991) randomized 63 patients to receive 20 weeks of group CBT, group experiential therapy, or a self-directed form of bibliotherapy (patients were asked to read a number of self-help texts, none of which were based on CBT or experiential models; each patient was contacted for about 30 minutes a week by a researcher (indicating that this would be better described as a minimal contact therapy). At 3-month follow-up, all three treatments had equivalent outcomes.

Computer-Aided Therapy

Selmi et al. (1990) solicited 36 patients with mild or moderate depression through newspaper announcements, randomizing them to one of three conditions. The two active conditions both involved a 6-week structured CBT intervention, delivered either by a therapist or through an interactive computer program; the remaining patients were allocated to wait-list control. At posttherapy and at 2-month follow-up, patients receiving both forms of CBT showed equivalent gains over those placed on the wait list; at follow-up, around two-thirds of those receiving CBT met criteria for remission, contrasted to only one wait-list patient. Although the active solicitation of patients limits generalizability, the majority of participants in this study had a depressive episode meeting research diagnostic criteria for more than 6 months, suggesting that their difficulties did not reflect transient distress.

In a small trial, Bowers et al. (1993) allocated 22 depressed inpatients to 2 weeks of CBT delivered by a therapist, 2 weeks of a computer-based therapy,

or to treatment as usual. At discharge, those receiving the computerized therapy were unimproved. The computer program has been criticized for relying only on cognitive rather than behavioral techniques (Marks, 1999), though whether this is the basis for differential outcomes is not clear.

A large-scale trial (Proudfoot et al., 2003) randomized 167 patients to receive an eight-session, computer-based CBT program for depression, or TAU from their primary care physician. The package was standardized but became customized to patients as they interacted with the program. Inclusion criteria did not restrict the sample to those with depression alone—nearly half were diagnosed with mixed anxiety and depression. Recruitment was by both referral from a primary care physician and screening in general practitioners' waiting rooms, using the General Health Questionnaire (GHQ-12). (The broad inclusion criteria and the use of screening [which would lead to patients entering the trial who may not have otherwise come forward for help] raises at least some question about contrast of results from this to other trials of treatments for depression.)

Though patients receiving computer-aided therapy received pharmacotherapy and any nonspecific support associated with TAU, they were precluded from receiving any form of concurrent psychological therapy. In addition, the treatments received by all patients were monitored. At posttherapy and through to 6-month follow-up, patients receiving the CBT program compared to TAU evidenced a significant reduction in BDI scores; mean pretreatment scores for CBT and TAU, respectively, were 25.5 and 24.0, and at 6-month follow-up, 9.5 and 16.0. Post hoc analysis suggested that CBT patients evidenced equivalent gains whether or not they were in receipt of medication, and that while initial severity did not predict outcome, duration of illness did: Patients who had been depressed for more than 6 months prior to trial entry showed significantly greater impairment at assessment points. This study provides a helpful pointer to the potential utility of computer-aided packages but is at present unique and requires replication.

Summary

Although a number of recent studies add to our knowledge of the relative impacts of pharmacotherapy and psychotherapy, there is continuing variability in outcomes from individual trials. Nonetheless, there is good evidence of the superiority of either active treatment over placebo. In relation to each modality offered alone, psychotherapy does not appear superior in efficacy to medication; only rarely does medication show clear advantage over psychotherapy and, on the whole, a reasonable conclusion would be that psychotherapy is of equivalent efficacy to medication. While there is some suggestion of benefit to combination treatment, this is hard to demonstrate. It is possible that the impact of combination treatment (which is widely practiced

in routine settings) is less in its additive impact than in facilitating the acceptability of treatment—for example, better accommodating issues, such as the slower speed of change for psychotherapeutic interventions with more severely depressed individuals, or the attrition associated with the impact of side effects in pharmacotherapy.

Although most medication–psychotherapy contrasts are to CBT, there is some diversity in terms of psychotherapeutic approach studied. As yet, there are not enough trials to detect whether there are variations in the way specific therapies interact or contrast with pharmacological approaches, though, on the basis of limited evidence, patterns of outcome seem similar.

Though rather limited, information on the relative acceptability of medication and psychotherapy suggests that psychotherapy or combination treatment is associated with lower attrition rates than pharmacotherapy alone. Definitive conclusions about pharmaco- and psychotherapies are hampered by methodological issues of both a specific and a conceptual nature. A lack of rigor and therapeutic equipoise in earlier trials almost certainly acted to enhance the relative benefits of psychotherapy, while problems with blinding may have led to overestimation of medication effects. More recent trials seem better designed and, hence, less susceptible to basic problems of interpretation.

Brief dynamic therapy, while widely practiced, continues to have a limited evidence base. Where good-quality trials have been conducted, outcomes are equivalent to alternative psychotherapies, but the paucity of trials represents a serious concern and severely limits conclusions that can be drawn about this approach. This position contrasts the increasing evidence base for a range of approaches—in particular CBT, IPT, couple therapy, problem-solving therapy, bibliotherapy, and computer-aided therapy, all of which show evidence of efficacy, though not necessarily of advantage relative to other approaches.

A pertinent observation is that all structured psychotherapeutic approaches show short-term efficacy in around 50–60% of cases. While this could reflect the impacts of specific elements of each therapy, it also raises the possibility that nonspecific responsiveness of many patient samples, together with regression to the mean, could account for this equivalence of action. In this respect, the concerns expressed by Klein and others (discussed earlier) may have some force. This does not imply accepting all their arguments, but without greater control for nonspecific effects, it is hard to be certain that short-term outcomes can be attributed to the technical ingredients of therapies. It is possible that an alternative explanation for the consistency of outcome is that this reflects a maximum treatment effect imposed by the nature of research samples and the lengths of treatment. If this were so, the important difference between treatments would be sought not in immediate treatment effects but in their capacity to delay relapse, a measure of outcome that more closely maps to the chronic nature of depression in many individuals.

MANAGING RELAPSE AND RECURRENCE

Though predicting relapse in individual patients is difficult, there is consistent evidence that the probability of experiencing a further episode of depression is greater if patients have achieved only partial remission after treatment. A number of trials (e.g., Paykel et al., 1995; Thase et al., 1992; Van Londen et al., 1988) reported that the presence of residual symptoms was associated both with higher levels of relapse and with shorter intervals between the previous and subsequent episodes. Residual symptoms were associated with more severe initial illness and with a history of previous depressive episodes but not with other indicators such as duration of depression or the presence of dysthymic disorder. Contrasted to patients who were asymptomatic after treatment, relapse appears to be about three times as likely in patients with residual symptoms, and to occur about three times more quickly (Judd et al., 1998a, 2000).

A claim often made for CBT is that it acts prophylactically against recurrence of depression, in part because the therapy addresses the dysfunctional cognitions thought to contribute to depressed states. There is little evidence of such specificity of action. For example Gortner et al. (1998) contrasted the impact of three treatments: (1) exclusively cognitive, (2) exclusively behavioral, and (3) both elements. Posttherapy data on 137 patients suggested that all three treatments were of equal efficacy, and relapse rates and survival time to relapse at 2-year follow-up were equivalent. Although there are now a number of trials suggesting that CBT lowers relapse rates, the contrast is often against no treatment or against patients withdrawn from medication. Since there is little research into the potential efficacy of other forms of psychological therapy, it is not clear that these reduced relapse rates are specific to CBT. Finally, though a reduction in relapse can be demonstrated, this is better seen as relative rather than absolute, since a significant number of patients suffered a further episode of depression.

Kovacs et al. (1981) followed up patients in the Rush et al. (1977) study; after 1 year, treatment gains with both CBT and medication were maintained, though patients treated with CBT had significantly lower levels of depression than those treated with medication, and there was a trend for more of them to be judged as being in remission. Beck et al. (1985) found improvement to be stable over 6- and 12-month follow-up in patients who had received CBT, or CBT in combination with medication. There was a trend for the combination group to do better than the CBT group at 12 months. Simons et al. (1986), reporting follow-up data from a study by Murphy et al. (1984), found that at 12 months, patients who had received CBT had a significantly lower relapse rate than those receiving medication. Blackburn et al. (1986), reporting follow-up data of Blackburn et al. (1981), found significantly greater relapse in patients who received medication than

in patients receiving CBT at 6 months, though this reduced to a trend at 2 years. Rotzer-Zimmer et al. (1985, cited in Williams, 1992) and Evans et al. (1992) also found a significantly reduced rate of relapse in patients treated with CBT compared to those receiving medication. In contrast, the follow-up phase of the NIMH study (Shea et al., 1992a) suggested no significant advantage to CBT over other interventions.

Though there is evidence for (at least a limited) prophylactic effect of short-term CBT, the evidence from most of the trials reviewed above is clear: Individuals with MDD are at risk of relapse following circumscribed interventions. Despite this, rather few trials examine the impact of adding a continuation phase to short-term interventions. Frank et al.'s (1990) seminal study was described in detail earlier; briefly, this study found some benefit to monthly IPT over 5 years for individuals with recurrent depression, defined as three or more episodes of MDD, with the most recent being no more than 2½ years prior to the present episode.

Blackburn and Moore (1997; reviewed earlier) followed patients over 2 years, though attrition makes it inappropriate to consider data beyond 12 months. All patients were defined as suffering "recurrent" depression, though the criterion for "recurrent" was set at one previous episode of MDD (a somewhat less rigorous marker than that used by Frank et al. [1990]). All received active maintenance treatment (either medication or monthly CBT); both treatments showed equal efficacy in terms of subsequent relapse. Of the 49 patients in this trial initially treated with medication, 13 failed to respond at the end of the acute phase of treatment. Seven of these patients were randomized to continue on medication, and six, to receive CBT. In the continuation phase, attrition further reduced this already small sample (Moore & Blackburn, 1997), and though more of those receiving CBT were categorized as showing a full or partial response than those receiving medication, the difference was not statistically significant.

Fava et al. (1998b) reported a seminaturalistic design that followed the progress of 40 patients with recurrent depression (defined using Frank et al.'s [1990] criterion) treated between 3 and 5 months with a range of antidepressant medications. At this point, patients were randomized to receive ten 30-minute sessions of CBT or clinical management, and medication was tapered and discontinued. Immediately after this phase, patients with CBT showed a significantly lower level of residual symptoms, and at 2-year follow-up had a significantly lower relapse rate (25% contrasted to 80% for clinical management). It should be noted that the therapy applied was a modification of standard CBT (though it retained recognizable elements of this approach) conducted by one therapist for all patients.

Fava et al. (1998a) reported outcomes over 6 years following an initial phase of treatment (Fava et al., 1994), in which 40 patients with MDD who had been successfully treated with antidepressants were randomly assigned to

10 sessions of either fortnightly CBT or clinical management; in both groups, medication was tapered and discontinued. Over the length of follow-up, relapse rates in the CBT groups were lower than those in clinical management, though this difference was only statistically significant at 4-year follow-up (at 2 years, respective relapse rates were 15% and 35%; at 4 years, 35% and 70%; and at 6 years, 50% and 75%).

Paykel et al. (1999) recruited 158 patients whose MDD had only partially remitted following an initial phase of treatment with appropriate levels of antidepressant medication. Patients were then assigned either to clinical management alone, or to clinical management combined with 16 sessions of cognitive therapy over 20 weeks (with two further booster sessions 6 and 14 weeks later). Unlike Fava et al. (1994, 1998a) they continued to prescribe medication throughout this phase. Patients in receipt of cognitive therapy showed a significantly reduced cumulative relapse rate. Based on the intent-to-treat sample, at 68-week follow-up, relapse rates were 29% contrasted to 47% for controls. Though relatively few patients met criteria for remission (stable subclinical scores on the HDRS and BDI over 4 weeks), significantly more experimental than control subjects achieved this status (at 20 weeks, 25% and 13%, respectively). Scott et al. (2003) reported a cost–benefit analysis of this study, which suggests that the additional cost of providing CBT was somewhere between £4,000 and £5,000 per relapse prevented (despite evidence of a marked reduction of in- and day-patient services). In this respect, it can be concluded that the addition of CBT is more costly but also more effective (particularly if it impacts on the longer term course of the disorder).

In contrast to these studies, Jarrett et al. (1998, 2001) did not include pharmacotherapy in any arm of their trial. On the basis of earlier promising outcomes from a nonrandomized pilot trial, 156 patients were entered into an acute phase of cognitive therapy comprising 20 sessions over 12–14 weeks. After this initial intervention, 84 patients agreed to be randomized either to receive 10 further sessions of therapy over 8 months or to a control condition in which they were monitored but no further intervention was offered. At this point, relapse rates for patients in receipt of continued therapy were significantly lower than in controls (10% and 31%, respectively). Stratifying the sample in relation to age of onset (before or after age 18) suggested that earlier onset was associated with greater vulnerability to relapse; 67% of control patients with onset prior to age 18 relapsed over 8 months contrasted to 36% of those with later onset. Though continuation therapy was especially beneficial for "early onset" patients, with significant reduction in relapse rates compared to controls (16% and 67%, respectively), it showed less impact for patients with later onset (relapse rates of 50% and 36%, respectively). This pattern was maintained at 24-month follow-up. A further subanalysis confirmed that individuals who failed to achieve stable remission in the later

phases of acute treatment were also vulnerable to relapse. Contrasted to controls, at 24 months, relapse rates for patients with unstable remission were significantly reduced by continuation therapy (62% and 37%, respectively).

Mindfulness-based cognitive therapy (MBCT) has developed as a way of conceptualizing vulnerability both to depression and to relapse in patients with a history of MDD (Teasdale, 1999). The conventional CBT model assumes that vulnerability to depression arises from dysfunctional beliefs or attitudes, which are modified by successful intervention. MBCT assumes that it is not so much the dysfunctional content of beliefs or attitudes that lead to depressive states as the facility with which patterns of negative thinking can become activated when individuals become dysphoric, and the ease with which these can rapidly escalate into a ruminative cycle. Vulnerability to depression arises not only because of the accessibility of negative thoughts but also because individuals find it difficult to gain a sense of perspective from which to appraise themselves. In practice, MBCT aims both to help patients increase awareness of their patterns of thought and to foster their capacity to appraise their cognitions from a "decentered" or disidentified position. On this basis, there is no direct challenge to cognitions (as in conventional CBT), but an attempt to alter the degree to which individuals react to these cognitions as if they were isomorphic with their sense of self.

Teasdale et al. (2000) explored the utility of MBCT in a multicenter trial based at three treatment sites (in the United Kingdom and Canada). Inclusion criteria required patients to have a history of two or more episodes of MDD within the past 5 years, with at least one of these episodes in the 2 years prior to the study. One hundred forty-five patients in recovery or remission from MDD were randomly assigned either to MBCT (delivered in a group format over nine weekly sessions) or to TAU. Analysis of relapse patterns over the 60 weeks of the study period suggested different outcomes for patients with three or more previous episodes of depression (who constituted 77% of the intention-to-treat sample), compared to those with two previous episodes. In the former group, significantly fewer of those in receipt of MBCT relapsed contrasted to those given TAU (66% and 40%, respectively). Among patients with two prior episodes of depression, relapse rates across treatment conditions were statistically equivalent (56% of MBCT and 31% of TAU group). Further analysis suggested that for TAU, there was a linear relationship between the number of prior episodes and the risk of relapse (31% for patients with two prior episodes, 56% for three, and 72% for four). This pattern was not present for patients in receipt of MBCT. Subanalyses suggested that the reduction in relapse rate was not attributable to increased medication prescription (which was comparable across treatment conditions). Patients with more than three episodes of depression experienced their first episode of depression at a significantly younger age than those with two prior episodes. This observation suggests the utility of identifying a subgroup of especially

vulnerable individuals for whom MBCT might be especially beneficial, and echoes Jarrett et al.'s (2001) suggestion that age of onset may be an important marker of vulnerability. Furthermore, both trials only demonstrated differential effects on relapse for individuals with a significant history of depression; if replicated by other trials, this would be a helpful marker for clinical intervention.

Summary

Medication, CBT, and IPT have shown efficacy in reducing relapse in patient samples selected on the basis of their vulnerability to relapse, usually using "booster" or maintenance sessions. There is evidence of the efficacy of maintenance psychotherapies combined with medication and also when offered alone in the context of discontinuation of medication.

Provision of maintenance sessions appears to be effective, but it is also costly. In view of this, it is helpful to have some preliminary indications of differential benefit for individuals with early onset and a history of previous relapses. This finding does not suggest that maintenance therapies should be restricted to this group; though they might be privileged for such an intervention, their higher response rate presumably rests on their known excess vulnerability to relapse.

Whether some therapies are better than others at reducing relapse is, as yet, a moot point. While it is encouraging to see the development of therapies (such as MBCT) that have developed from theoretical ideas about the nature of relapse, their benefit over standard technique is, as yet, unclear.

STUDIES OF EFFICACY IN DIFFERENT TREATMENT CONTEXTS

Inpatient Treatment

Although many trials of CBT with inpatients have been reported, five studies are of particular interest in that they examine the use of CBT with patients with more severe depression, and with associated behavior likely to exclude them from other treatment trials, such as suicidal behavior.

Thase et al. (1991) treated 16 unmedicated inpatients characterized by an HRSD score of ≥ 15 and with an index episode of MDD of less than 2 years' duration. All were drug-free for at least 7 days before the trial commenced. Twenty-six patients were assessed as suitable for the trial; 16 of these completed treatment, while the remainder had electroconvulsive therapy (ECT) or medication, either because of noncompliance with therapy or the emergence of severe symptomatology before therapy commenced. Intensive CBT was offered five times a week over 4 weeks; on average patients received 13 sessions of therapy. Response was defined by reduction in HRSD scores of at

least 50% and a final score of ≤ 10, and 13 patients (81%) reached this criterion. Follow-up therapy was offered, but only seven patients received more than 1 month of outpatient CBT; though the follow-up period is not specified, Thase et al. (1991) report that of these patients, only one relapsed, compared to three out of four patients who refused further therapy and whose progress was monitored.

Thase and colleagues (1993) reported an extension of this work to larger samples in three research trials (Nofzinger et al., 1993; Simons & Thase, 1992; Thase et al., 1991). In total, 142 unmedicated patients were treated either as outpatients ($n = 110$) or inpatients ($n = 32$). Outpatients received up to 20 sessions of CBT over 20 weeks; inpatients received more intensive therapy—20 sessions over 4 weeks (as reported in detail earlier [Thase et al., 1991]). Across all three patient samples, significant reductions in HRSD scores were found, though higher initial levels of depression were associated with poorer response rates. This effect was most marked for patients with HRSD scores above 20.

Bowers (1990) conducted a comparative trial of nortriptyline alone, relaxation in combination with nortriptyline, or CBT and nortriptyline, offered to 30 inpatients in addition to the usual hospital milieu. Therapy was conducted in groups, and 12 therapy sessions were offered. Forty-one patients were approached, eight declined, and one patient per group dropped out "because of violation of the protocol."

Patients were moderately to severely depressed; the mean pretreatment BDI scores for the CBT, relaxation, and medication groups were 24.2, 25.8, and 31.2, respectively (giving a nonsignificant trend toward greater initial severity in the medication group). All therapies were offered by the same therapist.

Symptoms were assessed using the BDI, the HRSD, and measures of cognitive adjustment at sessions 1, 6, and 12, and at discharge. All groups improved, but patients receiving CBT or relaxation had significantly fewer depressive symptoms and negative cognitions than patients in the medication-alone condition. In addition, patients receiving CBT were less likely to be judged depressed at discharge than those in the other treatment conditions. A recovery criterion of an HRSD score of ≤ 6 was achieved by 8 of 10 patients in the CBT group, compared with 1 of 10 and 2 of 10 patients receiving relaxation or medication alone.

The degree to which this result reflects the specific impact of CBT is not clear. Using a criterion based on a BDI score of ≤ 9, patients receiving relaxation showed similar gains to those in the CBT group. Interpretation of this study is also made more difficult by the fact that there was no control for the additional attention psychotherapy patients received in contrast to those on medication alone.

Miller et al. (1989) assigned 47 patients to one of three conditions—standard treatment (hospital milieu, medication, and medication manage-

ment), CBT and standard treatment, or social skills training and standard treatment. All patients had BDI scores ≥ 17 and HRSD scores ≥ 17. Therapies were conducted daily while the patients were in the hospital, and continued weekly after discharge. CBT was offered daily while patients were in the hospital and weekly after discharge. All therapies led to significant gains on a range of measures. At discharge, there was a trend for patients receiving combination treatments to be categorized as responders; after outpatient treatment, the trend reached significance. However, there were significant differences in the dropout rate between conditions—41% from standard treatment, 31% from CBT, and 14% from social skills. In addition, all patients from the standard treatment group had dropped out by week 8 of follow-up, leading to problems in interpreting the follow-up data.

Scott and colleagues in Newcastle present data from two open trials of combined medication and CBT, offered to chronically depressed inpatients who had previously failed to respond to standard antidepressants and had been depressed for at least 2 years. Pharmacotherapy comprised phenelzine, L-tryptophan, and lithium. In the first trial (Barker et al., 1987), 20 patients were randomly assigned either to pharmacotherapy alone or to combination treatment with CBT (delivered biweekly for 3 weeks, followed by nine weekly sessions). Though 11 patients showed a 50% reduction in HRSD scores (all within the first 6 weeks), there was no evidence for an additional benefit from CBT. In a second trial with a similar population (Scott & Freeman, 1992), 24 patients were divided into two cohorts. The first (n = 8) received 12 weeks of combined pharmacotherapy and CBT as described in Barker et al. (1987), with similar outcomes to those described earlier. The second (n = 16) was offered a modified CBT package with a "milieu" treatment; patients were admitted to a dedicated inpatient unit, and therapy was more intensive and prolonged—approximately 26 inpatient sessions followed by at least 6 months of outpatient treatment. Percentage change scores on the BDI and HRSD were greater for patients receiving the modified package (52% and 57%, respectively) than for those receiving standard CBT (42% on both measures), and significant change was observed in 69% of patients. Although suggestive, small sample sizes and nonrandom allocation limit the conclusions that can be drawn.

Treatment in Primary Care

Most research focuses on patients referred to specialist services, and though outcomes from this work are almost certainly applicable to those seen in primary care, it is not safe to assume that the same outcomes will be achieved. Patients seen by primary care physicians probably represent a broader range of presentations than those seen in secondary care, and the clinical picture is often complicated by somatic presentations. For many patients, their first (and

sometimes only) port of call is primary care; on this basis, this section considers whether specific interventions have greater efficacy than TAU offered by primary care physicians. TAU is a nonspecific comparison, potentially containing a number of uncontrolled elements—not only interventions offered by the physician but also treatments offered by specialists to whom the patient is referred. This means that TAU can vary significantly in relation to the treatments offered to individual patients, and in relation to the practices of different family physicians, reducing internal validity. In addition, any differences that emerge between TAU and a comparator treatment will reflect the quality of local primary care services and the services to which it has access. Although contrast to TAU is ecologically appropriate, care is needed in interpreting results from these studies.

Teasdale et al. (1984) treated 17 patients with BDI scores ≥ 20, contrasting them with 20 patients receiving TAU. Although CBT led to a significant difference in the number of patients judged recovered posttreatment (indicated by a BDI score ≤ 10), at 3-month follow-up the TAU group had also improved, leading to no between-group differences at this point. Ross and Scott (1985) treated 51 patients with BDI scores ≥ 14; patients continued to receive TAU from their family physician but were additionally allocated either to individual or group CBT, or to a 3-month wait-list control group (that subsequently received CBT). A 64% reduction in BDI scores was found for the CBT group, contrasted to a 13% reduction in the wait-list group. However, no figures using a recovery criterion are given in the study, and relapse was defined as a BDI score ≥ 16, which is markedly less stringent than that usually adopted. Partial data from a 12-month follow-up suggested that no patients receiving CBT relapsed on this criterion.

Scott and Freeman (1992) requested that 63 family physicians from 14 primary health care practices refer patients with a depressive disorder. One hundred ninety-four patients were referred and 121 were accepted into the trial. The study design was such that some patients would be assigned to treatment with medication; of some interest is the fact that most patients who declined to take part in the study cited as a reason a reluctance to take medication. Patients were randomly assigned to one of four conditions for 16 weeks of treatment. Help was offered by a psychiatrist (for amitriptyline), a clinical psychologist (for CBT), or a social worker (for supportive counseling). In the remaining condition, patients were reassigned to their family physicians for TAU. One difficulty in this study is that randomization of patients to treatment conditions was not successful; only 11 of 29 (38%) of patients seeing the social worker had HRSD scores ≥ 16, suggesting that most clients in this group did not achieve a level of "caseness" for depression (contrasted to 22 of 30 [73%] patients in the family physician group). In addition, only 2 of 29 (7%) patients seeing the clinical psychologist had a previous episode of depression. At the end of treatment, only social work counseling

showed a greater reduction in depressive symptoms when contrasted to care from the physician. It has already been noted, however, that patients in each of these groups differed markedly in their initial levels of depression, which makes it difficult to interpret this result. Patient satisfaction was greatest with social work counseling, though the fact that only one therapist offered each treatment modality increases the likelihood of therapist-specific effects.

Schulberg et al. (1996) randomized 276 patients with a diagnosis of MDD to receive nortriptyline, IPT, or TAU. A very large number of patients (7,652) attending primary care services were screened for depression, and further filters ensured that all patients met DSM-III criteria for depression and had an HRSD score ≥ 13. For both active treatments an acute phase determined whether patients were treatment responders; for medication, this phase continued over 6 weeks, and for IPT, over 16 weeks, with improvement defined as a 33–50% reduction in initial BDI score. Those who met these criteria entered a 4-month continuation phase (with sessions at monthly intervals). There was significant attrition from both phases of the trial. Only 50% of patients completed the acute phase; in the continuation phase, 40% of patients receiving medication and 20% of those in receipt of IPT dropped out. Though symptom levels in the intent-to-treat sample reduced across all interventions, the active treatments showed equal efficacy, and both resulted in significantly greater gains over TAU. A similar pattern was evident for patients who completed the continuation phase. At 8 months, a recovery criterion of an HRSD score ≤ 7 was met by 48% and 46%, of patients receiving medication and IPT, and by 18% of those receiving TAU.

Scott et al. (1997) contrasted the efficacy of brief CBT plus TAU to TAU alone. Forty-eight patients with a BDI score ≥ 20 and a depressive episode of less than 2 years were randomized to treatment. CBT was delivered in six sessions and followed a systematic but flexible protocol that was adapted to each patient; all therapies were offered by the same therapist. At 7 weeks, BDI scores in the CBT group were significantly lower than those for patients receiving TAU. Although data from the follow-up period suggest greater gains for CBT at 1 year, significantly greater attrition from TAU, combined with a low sample size, suggests that this result should be interpreted cautiously.

Corney (1987) contrasted 80 depressed women receiving either routine treatment from their family physician or social work counseling. No clear model of counseling was followed, though counselors reported using exploration, practical help, and some behavioral goal setting. Overall, there was little difference in outcome among treatment groups. The sample was stratified according to the degree of severity of depression and its chronicity. Patients with more acute and less severe problems improved regardless of treatment received, with more moderate outcomes in more severely distressed patients. There were some indications that, contrasted with equivalent controls, those

patients with acute but more severe problems had better outcomes when in receipt of counseling.

Raphael (1979) examined the efficacy of counseling for bereaved women considered to be at risk of delayed or pathological grief reactions. In an initial pool 200 women were interviewed. On this basis, 64 patients were selected who demonstrated either marked ambivalence in their relationship to their husbands and/or had poor social support for their grieving. The subgroup of 64 patients was randomly allocated either to counseling or no-treatment groups. Counseling was based on psychodynamic/exploratory methods, focused on the bereavement, and was offered for the 3 months following the death. At 13-month follow-up, 77% of the counseled group had good outcomes, contrasted with 41% of controls.

Holden et al. (1989) reported a trial of counseling for women with acute postnatal depression, delivered by health visitors who had been given a brief (3 week) course in nondirective methods. Forty-eight women were allocated either to eight weekly sessions of counseling, in addition to standard health visitor support, or to standard health visitor support alone. Approximately 3 months after treatment started, 69% of the counseled patients no longer met criteria for depression, contrasted with 38% of the control group.

It may be significant that in both Holden et al. (1989) and Raphael (1979), interventions were targeted, were specific to a client group with acute difficulties, and were delivered by therapists who would have been familiar with their patients' presentation. This draws attention to these specific characteristics of both counselors and patients, and may suggest that counseling interventions may be more likely to be successful when they are focal and focused.

Ward et al. (2000) undertook an ambitiously designed large-scale trial, aiming to examine the relative efficacy of TAU contrasted to 12 sessions of nondirective counseling or CBT, including a patient-preference arm, along with randomization. Four hundred sixty-four patients met entry criteria, which included a BDI score ≥ 14; only 62% of patients were diagnosed as depressed, and the sample is perhaps best characterized as mildly to moderately depressed. While randomization was encouraged, the choices of patients who expressed a strong treatment preference were honored. Monitoring of this strategy suggested that patients were reluctant to be randomized to TAU, though they had few preferences about the type of psychological therapy they received. Since this resulted in rapid recruitment to the two active treatments, a second tranche of patients was offered randomization to CBT or counseling (a two-way rather than a three-way randomization). Although—appropriately—results are reported in way that reflects this complex pattern of sampling, the overall pattern of outcomes for randomized and patient preference arms was not significantly different: At 4 months, both active treatments showed equivalent and significant advantage over TAU, though at 1 year, there were no differences in outcome.

Simpson et al. (2003) carried out a trial of psychodynamic counseling, identifying patients from a number of primary care sites partly through screening of attenders using the BDI, but also through referral from family physicians. On this basis, 143 patients were randomized to counseling or TAU; active treatment comprised 6–12 sessions of nonmanualized therapy. At 6- and 12-month follow-up, no differences between active treatment and control were evident, and while there was evidence that individuals with milder depression showed some benefit, this was not the case for patients with more severe depression (defined as an initial BDI score greater than 24). In addition, many of those who fell into this latter category continued to have scores classifying them as "cases" at follow-up. A parallel report (Simpson et al., 2000) suggests that results were similar for two additional counselors who used CBT rather than psychodynamic techniques. While a number of interpretations are possible, it is worth observing that the pattern of outcomes conforms to evidence from other trials indicating that more severely depressed individuals are unlikely to respond to therapy of this brevity (and that allocation of patients to therapies needs to attend to likely dose–response relationships).

Studies reviewed to this point contrast the impact of individual therapists delivering an intervention. Sherbourne et al. (2001) reported a different strategy, whereby primary care settings as a whole were randomized to receive one of two quality improvement (QI) programs, in which nurse specialists were trained to assess patients and to formulate a treatment plan. In the first (QI-meds), nurses informed patients that medication and therapy were of equal effectiveness, and offered them the option of receiving medication (delivered through the nurse) or counseling (offered in the context of the usual services available in the practice). The second program (QI-therapy), was directed by the family physician, who used the nurse's assessment to decide which patients were appropriate for psychological therapy; if they were, he/she referred them to local psychotherapists, who delivered 12–16 sessions of individual or group CBT, or (for patients whose symptoms were milder) a brief four-session CBT intervention. Medication was available from the primary care setting. The program was monitored every 6 months over 2 years, and the sample comprised 1,299 patients from a total of 27,332 consecutive attendees screened for depression and 3,918 patients identified as potentially eligible for the program.

Four hundred five patients received QI-meds, 464 received QI-therapy, and 430 received TAU. Of these, about half had a diagnosis of MDD or double depression, one-fourth, depressive symptoms in the context of a history of MDD, and one-fourth had depressive symptoms without a previous episode of depression. In the first and second 6 months of QI-meds, 51% and 43% of patients took medication; for QI-therapy, the equivalent medication rates were 39% and 35%, and 38% and 34% received at least four sessions of therapy.

Over the first year, patients receiving QI-meds or QI-therapy had a significantly reduced probability of depression contrasted to those receiving TAU. The trajectory of change was greater for patients in the two QI groups; at 6 months, the rate of probable depression was 51% for TAU, contrasted to 41% in both QI groups. At 1 year, depression rates in both QI groups remained stable, but the rate in TAU had declined (to 48%). The trajectory of this group was a slow improvement, with the result that, by the end of the second year, rates of depression in all groups were equivalent (at approximately 43%). Examining the trajectory of change over time for all three groups, there was evidence that patients receiving QI-therapy had a more stable pattern of change, and a lower probability of poor outcomes, than those receiving QI-meds or TAU.

Summary

There is good evidence that CBT, particularly when delivered intensively, can be a useful adjunct to treatment in inpatient settings and with more severe cases of depression. However, trials usually offer CBT in combination with other treatments, such as social skills training, bibliotherapy, and relaxation. Though this means that the specific effects of CBT are not well established, the effect of therapy is to reduce the severity of depression on discharge and follow-up, and lead to better levels of adjustment.

In primary care, CBT, IPT, and nondirective counseling all seem reasonably effective, though the usual contrast is to TAU, and any advantage is short term rather than long term. Outcomes suggest that interventions are transportable, in that what seems efficacious in secondary or tertiary care also seems to work in primary care. Although it might be expected that briefer treatments would be effective in primary care, the duration of therapies employed in many trials is similar to that in more specialist contexts. While the efficacy of shorter and longer treatments appears equivalent, no trials directly contrast treatment length and, overall, there are too few trials to allow us to draw reliable conclusions.

TREATMENTS FOR DYSTHYMIC DISORDER

Prior to reclassification in DSM-III, patients with dysthymic disorder were seen as suffering a problem of personality rather than a mood disorder. However conceptualized, there is a clinical need to understand the most effective way of managing such patients. Though their depressive symptomatology may be mild (making them appear less needy), the chronicity of their condition appears to result in greater social incapacity than that experienced by many individuals with MDD (Klein & Hayden, 2000; Klein et al., 2000). Furthermore, they are at enhanced risk of suffering an episode of MDD (at

which point their condition would be described as "double depression"). Many present in a primary care context, but until recently, there has been little substantive research into the management of individuals who present as dysthmymic (as contrasted to the treatment of such individuals once they enter into a depressive episode).

Markowitz's (1994, 1996) reviews of psychological therapy for dysthymic disorder noted that, at this point, there were no available psychodynamic studies, some small-scale open trials of IPT (Markowitz, 1994; Mason et al., 1993), and seven open trials of CBT. Subsequently, a number of controlled trials have been published, though most contrast the relative impacts of medication and psychotherapy in various combinations. The absence of studies of the efficacy of psychotherapy alone is striking.

Cognitive-Behavioral Therapy

Gonzales et al. (1985) treated 113 patients with 12 two-hour individual or group "psychoeducational" sessions over 2 months, with follow-up at 1 month and at 6 months. Results varied according to diagnosis, with more improvement for those with acute MDD (75% reaching a recovery criterion) than for those with chronic intermittent depression (43%) or double depression (27%). De Jong et al. (1986) treated 30 unmedicated inpatients over 3 months. A combination of activity scheduling, social competence training, and cognitive restructuring achieved a higher response rate (60%) than cognitive restructuring alone (30%) or a wait list (10%). However, data from dropouts from treatment were not analyzed. At 6-month follow-up of half the sample, gains were maintained. One problem with this study is that response was defined as a BDI score ≤ 14, or as a 50% reduction in pretreatment BDI scores. The clinical significance of gains defined in this way is arguable.

Five very small-scale studies were identified by Markowitz (1994); in all cases, the sample size renders them exploratory. Fennel and Teasdale (1982) treated five patients with long-term depression, all of whom had failed to respond to previous treatment; only one patient showed clear improvement. Harpin et al. (1982) reported the treatment of 12 patients who failed to improve with medication. Patients either received 10 weeks of twice-weekly CBT ($n = 6$) or were allocated to a wait-list control group ($n = 6$). A significant drop in HRSD scores was found in the active treatment group as contrasted to the control group, though results were poorer with more severe levels of depression. Two of the six treated patients showed significant pre- to posttreatment improvement, but only one maintained this at 6 months. Stravynski et al. (1991) treated six patients with 15 weekly sessions of CBT; significant improvements in HRSD scores were obtained, and four patients no longer met criteria for dysthymic disorder following treatment. McCullough (1991) treated 10 patients with dsythymic disorder over a rather

longer period than the above-mentioned studies, with a range of 14–44 weekly sessions. All reached the recovery criterion of a BDI score ≤ 10, and nine remained in remission at 2-year follow-up. These results are perhaps less promising than they appear, in that in an original cohort of 20 patients treated, four did not complete treatment and six were unavailable to follow-up. Mercier et al. (1992) reported a 12- to 16-week trial with 15 patients with chronic dysthymic disorder; four booster sessions were offered over the 6-month follow-up period. Three of eight patients with dysthymica and three of seven patients with double depression responded, and of the six responders, four remained well over the follow-up period. Given that all responders had been depressed for 7 years or longer, this is an impressive result.

Contrasts of Psychological Therapy and Medication Alone and in Combination

Cognitive-Behavioral Therapy

Becker et al. (1987) allocated 39 patients either to social skills training or to crisis–supportive psychotherapy, along with either nortriptyline or placebo. After 16 weeks of treatment, gains were evident in all conditions. Dunner et al. (1996) contrasted the short-term treatment with CBT or fluoxetine in 24 patients randomized to 16 weeks of either treatment. Though at 8 and 16 weeks more patients receiving fluoxetine met criteria for recovery (7 of 13, contrasted to 2 of 11 receiving CBT), this was not statistically significant, and no follow-up data were collected.

Interpersonal Psychotherapy

de Mello et al. (2001) contrasted the impact of moclobemide alone and moclobemide combined with IPT. Thirty-five patients were randomized to each treatment; group therapy comprised 16 weekly sessions followed by monthly maintenance sessions over 6 months. Outcomes for both treatments were equivalent, though small initial sample sizes and consequent high levels of attrition limit interpretation of results from this trial.

Brown et al. (2002) conducted a large trial in primary care, allocating 707 patients to sertraline alone, to 10 sessions of IPT alone, or to the combination of these treatments. Patients were identified on the basis of epidemiological screening and advertisement. No information is given regarding therapist qualifications for conducting IPT, but random adherence checks were conducted (though it should be noted that the dosage of IPT is lower than in most trials). The acute phase of this trial took place over 6 months, with a further 18-month naturalistic follow-up. At posttherapy, sertraline alone and

combined treatment showed equivalent outcomes, and both were superior to IPT alone, with respective response rates of 60%, 58%, and 47%. At 2-year follow-up, there was evidence of lower health care utilization by patients in the combined group, but an obvious potential confound is that all patients were offered sertraline over follow-up, an offer taken up by about 60% of those receiving medication in the acute phase, but only 12% of those receiving IPT.

"Problem-Solving" Treatment

Problem-solving treatment has a psychosocial, here-and-now focus, and encourages patients to specify and work toward resolving areas of functioning that they identify as problematic. Most trials examine the impact of this approach for individuals with mild to moderate depression in primary care settings.

Mynors-Wallis et al. (1995, 2000) recruited patients through primary care physicians in Oxfordshire. They (1995) allocated 91 patients to one of three treatments—amitripyline and clinical management, placebo and clinical management, or six sessions of problem-solving treatment delivered over 12 weeks. Adopting a recovery criterion of an HRSD score ≤ 7 at 12 weeks, the recovery rate in patients receiving amitripyline and problem-solving therapy was equivalent (60% and 52%, respectively), but was significantly greater than for patients receiving placebo (27%). A later study by the same research group (2000) randomized 151 patients to receive problem-solving therapy alone, fluvoxamine or paroxetine alone, or problem-solving therapy combined with one or the other of these medications. Over 12 weeks, all treatments showed equivalent efficacy. It should be noted that entry criteria for these studies are lower than in most trials reported in this section (patients needed to meet research diagnostic criteria for depression and have an HRSD score ≥ 13—usually an indicator for mild depression). On this basis, it may be inappropriate to generalize these results beyond benefit to patients with mild to moderate depression.

Dowrick et al. (2000) report a multicenter, multinational study conducted in the United Kingdom, Ireland, Spain, Finland, and Norway, identifying suitable participants on the basis of a community survey. Four hundred fifty-two people were recruited (though it is unclear how this number differs from the potential pool of participants) and randomized to one of three conditions—six sessions of problem-solving therapy, one of two variants of group psychoeducation (over 8 weeks), or a no-treatment control. The design is complicated by the fact that only one site offered both interventions, that both rural and urban centers were included, and (relatedly) that problem-solving sessions were usually offered in the patient's home, whereas psychoeducation involved travel: Attrition for psychoeducation was signifi-

cantly greater than for problem solving (less than half the sample completed the course). At 6 months (though not at 12 months), the two active treatments showed gains over controls, with approximately 58% no longer meeting criteria for depression, contrasted to approximately 42% of controls. Though methodologically problematic, this study is unusual in describing the delivery of interventions in a community-based context.

Barrett et al. (2001) reported outcomes from a multicenter trial based in primary care, which randomized 241 patients with dysthymic disorder (127) or minor depression (114) to problem-solving therapy, paroxetine, or placebo. All patients had six scheduled treatment sessions over 11 weeks, with follow-up over 6 months. The criterion for remission was set at an HRSD score ≤ 6 at 11 weeks. For patients with minor depression, there was a high rate of remission (at around 64%) and no differences between interventions (though whether this speaks to the high level of contact for all patients or the tendency toward patient responsiveness is not clear). For patients with dysthymic disorder, there was a differential treatment effect: paroxetine and problem-solving therapy achieved a significantly higher remission rate than placebo (80%, 57%, and 44%, respectively). Though at 6 months no between-treatment differences were observable (Oxman et al., 2001), follow-up was complicated by the fact that (in effect) patients received treatment as usual posttherapy, making it inappropriate to attribute changes to the original randomization.

A second report from this group (Williams et al., 2000) followed the same research strategy as had Barrett et al. (2001) but focused on a sample of older adults (over age 60, with a mean age of 71); 211 patients presented with dysthymic disorder and 204 with minor depression. Again, problem-solving therapy was contrasted to paroxetine alone or to placebo over 11 weeks of treatment. On intent-to-treat analyses, patients with dysthymic disorder showed significant benefit from paroxetine but not from problem-solving therapy. However, marked site differences were found in remission rates for patients who received four or more sessions of either treatment; rates for problem-solving therapy ranged from 33 to 80%, and for paroxetine, from 27 to 67%. Though the authors note site-specific variations in therapist expertise in problem-solving therapy, there is no formal analysis of factors that might have contributed to this pattern of outcomes.

Group Therapy

Two trials have examined the impact of combining group therapy with medication for individuals with dysthymic disorder. Hellerstein et al. (2001) randomized 40 patients who had responded to 8 weeks of treatment with fluoxetine either to continue with medication alone, or to receive medication combined with 16 sessions of a manualized group therapy (which

included both a cognitive and interpersonal focus). An equivalent proportion of patients in each therapy met the recovery criterion, an HRSD score ≤ 7.

Ravindran et al. (1999) randomized 97 patients with dysthymic disorder to receive either sertraline alone, placebo alone, sertraline in combination with group CBT, or placebo in combination with group CBT, with group therapies conducted over a 12-week period. With response defined as an HRSD score ≤ 10 and at least a 50% decline in HRSD score at posttherapy, there was a higher (but nonsignificant) response rate for the combination of setraline and CBT contrasted to setraline alone (71% and 55%, respectively). There was also no difference in the response rate between CBT combined with placebo and placebo alone (33%), suggesting that, in this trial, CBT not only failed to enhance the benefits of medication but it also achieved no more benefit than placebo.

Summary

In recent years, there has been an increase in research focusing specifically on dysthymic disorder, with most of this increased attention contrasting the efficacy of psychological treatment against medication. Available contrasts suggest that adding psychological therapy to medication confers little advantage, and in some trials, medication alone showed greater efficacy than psychological therapy alone. Our earlier review of this area (Roth & Fonagy, 1996) was based on the small-scale trials extant at that point and tentatively concluded that there was some evidence of the efficacy of IPT and CBT. Since that time, there has been increased interest in problem-solving therapy but no equivalent focus on the use of therapies with proven benefit in more serious depressive disorders. In some ways, this is surprising; 79% of individuals with dysthymic disorder will eventually present with MDD (McCullough et al., 1992), and there is a natural course of remission and recurrence (Keller & Shapiro, 1982; Keller et al., 1983). After intervention for MDD, they are almost certainly at enhanced risk of subsequent relapse. From this vantage point, greater knowledge about the most effective management of patients with subthreshold symptoms would be an advantage; at present, the focus of researchers makes it harder to derive clear guidelines.

PROCESS FACTORS

Relating patterns of outcome to process factors is a difficult task. The effect sizes attributable to process factors are usually small, as are the sample sizes on which they are based. Only rarely do designs include process factors as main effects. Post hoc analysis can only look for associations between variables, but the success of this strategy is based on an assumption that these relationships

will be linear—an assumption that is almost certainly erroneous (e.g., Stiles & Shapiro, 1989; discussed further in Chapter 16). A further concern is that although analysis of any one data set may suggest process–outcome links, these may be findings that are specific only to the sample—a risk that is heightened by multiple analyses of the same data set. To some degree, the volume of research into depression offers the possibility of more robust research, but even here it is probably fair to say that there remain more questions about process factors than there are answers.

Therapeutic Alliance

As discussed in Chapter 16, most analyses find a significant association between measures of the therapeutic alliance and outcome. Krupnick et al. (1996) found that the quality of the therapeutic alliance (and especially measures of the patient-related alliance) predicted outcome in all arms of the NIMH trial, including pharmacological interventions. Castonguay et al. (1996) reexamined the University of Minnesota trial, taking measures of both the quality of the alliance and the degree to which therapists challenged dysfunctional assumptions (a core aspect of CBT technique). Importantly, they found that in the context of a positive alliance, greater cognitive challenge was associated with better outcomes, but that this technique exerted a negative impact if the alliance was negative. Stiles et al. (1998) found that, broadly speaking, a positive alliance was associated with positive outcomes in the Sheffield trial, though the detailed pattern of associations with particular aspects of the alliance was complex. Across a number of different therapeutic approaches, the influence of the alliance is well established. This raises important, but unresolved, questions about the interplay between the techniques embedded in "brand-name" therapies and the common therapeutic factors implied by the alliance concept.

Patient Characteristics and Outcome

Of these, initial severity, age at first onset, number of episodes, and chronicity at presentation have all been discussed earlier. In line with clinical observation, most (but not all) research suggests that each of these factors makes it more likely that patients will be less responsive to therapy, and this will be reflected by both a poorer outcome after short-term intervention and a greater probability of subsequent relapse. A combination of greater chronicity, severity, and earlier first onset tends to predict higher residual symptoms at the end of therapy (e.g., Agosti & Ocepek-Welikson, 1997), which in turn increases the risk of subsequent relapse (Hamilton & Dobson, 2002). Of course, this general observation may not apply in the individual case, and undue therapeutic pessimism may be inappropriate, but such find-

ings should at least alert clinicians to plan for likely patterns of treatment response.

Personality Disorder

Although there is a broad clinical consensus that the presence of a personality disorder leads to poorer outcomes (e.g., Department of Health, 2001), evidence to support this belief is not strong. Methodological issues are pertinent here. How personality disorder is defined, and perhaps more crucially, how it is measured, seems to impact on the degree of support for this assertion, and it appears that the better the quality of study design the less likely it is that a relationship will be found (Mulder, 2002). There is also the risk that assessments of personality disorder made while the patient is depressed may be unreliable (e.g., Stuart et al., 1992). Results from individual studies are not always consistent. In the NIMH trial, some 74% of patients received a diagnosis of personality disorder. Contrasted to individuals without personality disorder, mean depression scores were equivalent at termination, though there were poorer outcomes in relation to social functioning and a higher probability of residual symptoms—a pattern consistent across all clusters of personality disorder (Shea et al., 1992b). Analysis of outcomes for 27 patients who met criteria for DSM Cluster C (anxious–fearful) in the Sheffield trial (Hardy et al., 1995b) suggested that these patients had higher initial symptom levels but that those who received CBT improved to the same degree as patients without personality disorder. However, those treated with psychodynamic therapy had poorer outcomes. This later finding is echoed by two studies of psychodynamic treatments; both Diguer et al. (1993) and Hoffart and Martinsen (1993) reported that although individuals with personality disorder improved, they made smaller gains. Tyrer et al. (1993) reported on a cohort that included dysthymic (but not depressed) patients; they found that people with personality disorders tended to be less receptive to psychological therapies (in their trial, CBT), but were more responsive to medication. Kuyken et al. (2001) conducted a post hoc analysis of 162 depressed patients treated under naturalistic conditions. Fifty-nine percent of this sample was diagnosed with a personality disorder; though their initial symptom levels were higher, outcomes were equivalent to those without this comorbidity. Despite this equivalence of action, there was some evidence that some types of beliefs linked to personality disorder—specifically, avoidant and paranoid beliefs—were associated with poorer outcome. It is hard to discern a definitive pattern in these results, though more structured therapies appear to have greater capacity to produce an equivalence of action in the face of comorbidity. It is also relevant that differential impact appears not to be located in the domain of symptomatic change, but in relation to aspects of presentation that intertwine with the notion of personality disorder itself, such as interpersonal

functioning and the level of background symptomatology both pre- and posttherapy.

Personality Type

Barber and Muenz (1996) suggest that avoidant patients will be more responsive to CBT (because it will encourage them to confront feared situations), while obsessive (hence, unexpressive) patients will benefit from more expressive therapies (such as IPT) that facilitate emotional expression. Though 32 completer patients met DSM-III criteria for these two categories, sample sizes across the four "arms" of the analysis were inevitably low. Nonetheless, their results were intriguing, complicated by the fact that nonmarried patients had better outcomes with IPT, and married patients, with CBT (a result that may make sense in the context of IPT's focus on interpersonal gains, and that there may be less potential for gains in this area for married patients). Holding this factor constant, the predicted relationship between personality type and outcome was found. A somewhat different picture emerges from the Sheffield trial (Hardy et al., 2001), in which patients who tended to distance themselves from relationships (in this sense, avoidant) did less well in CBT than those who were more interpersonally engaged, though the presence of a positive alliance mitigated the impact of this factor.

Perfectionism

Sotksy et al. (1991) found that, overall, higher levels of perfectionism were associated with poorer outcomes. Blatt et al. (1996) found that in both CBT and IPT, outcomes for patients low and high in perfectionism were effectively unrelated to the development of the alliance. It may be that those low in perfectionism are able to tolerate therapeutic imperfections, while those high in perfectionism are relatively impervious to anything other than a negative view of self and other. However, outcomes for patients in the midrange of perfection were significantly related to the strength of the alliance, suggesting that these patients will be sensitive and potentially responsive to variations in therapist style. There are indications that these patients do not show the usual pattern of an increasing engagement as therapy progresses (Zuroff et al., 2000), suggesting that therapists may need to be focused on strategies to enhance patients' capacity to be active collaborators in their own therapy.

Summary

Although most trials are organized in relation to interventions, there is evidence that a positive therapeutic alliance is associated with better outcomes. The temporal relationship between alliance and symptomatic improvement is

probably rather complex (a matter discussed further in Chapter 16). However, it seems reasonable to suggest that in the absence of a positive alliance, specific technical interventions are unlikely to be effective.

Patients who present the greatest therapeutic challenge can be described fairly clearly: They are more likely to have had an early onset of depression and many previous episodes. It needs to be borne in mind that while these statistical associations are reasonably consistent across studies, there is considerable variability in the response of individual patients, even within studies. Importantly, severity of depression is not necessarily a negative indicator, presumably because the intensity of symptoms may be a poor guide to their underlying determinants.

While there is reasonably consistent evidence that personality disorder (more than personality type) impacts negatively on outcome, it should not be assumed that the presence of Axis II comorbidity necessarily attenuates treatment effects, in part because it is conceptually and methodologically difficult to disentangle deficits that relate to personality disorder from those attributable to depression. Persons with personality disorder may make therapist adherence to technique more difficult and, hence, require therapists to be more competent in their delivery of technique and their management of the alliance. In this sense, variations in findings may reflect the capacity of trialists to apply the treatment protocol—a particular test of alliance–technique interactions.

SUMMARY AND CLINICAL IMPLICATIONS

Depression is both common and chronic; reflecting this, a large body of research addresses its treatment. DSM-IV-TR makes an important distinction between patients presenting with acute episodes of depression and those who suffer from depression in a less intense but more chronic form (dysthymic disorder). Although (at least potentially) the latter represent a greater clinical challenge than the former, relatively few studies focus on their treatment.

Evidence from meta-analytic review combined with consideration of individual studies demonstrates short-term efficacy for structured psychological therapies offered in brief formats (usually around 16 weeks), with no clear evidence of advantage to any particular approach. Although these initial gains are clinically significant, rather few trials have extended follow-up, and where this exists, there is a clear tendency for patients to relapse. A number of studies (using IPT or variants of CBT) have investigated the benefits of maintenance or "booster" therapy, and there is evidence for the efficacy of this approach in reducing the recurrence of depression.

Although there is an increasing amount of research on dysthymic disorder, nearly all major trials focus on the adjunctive use of psychological ther-

apy and medication, with little evidence that psychotherapy either adds benefit to pharmacotherapy or is effective when offered alone. The inappropriate characterization of dysthymic disorder as a minor form of depression has probably contributed to this lack of research. This is unfortunate, partly because we know less than we should about how therapies commonly used to treat MDD would perform if applied to dysthymic disorder. Equally, knowing more about how best to manage individuals with subthreshold and chronic depressive symptoms would be helpful when considering the treatment of MDD, since we know that individuals who continue to have subthreshold symptoms after treatment are at elevated risk of relapse.

Trials contrasting pharmacotherapy and psychotherapy suggest an equivalence of action between the two approaches, and it has been difficult to demonstrate that their combination is more effective than either offered alone (though studies conducted in inpatient contexts and with dysthymic disorder may be an exception to this conclusion). The reliability of studies contributing to this comparison is not as robust as we might wish, both in terms of the quality of pharmacotherapy and the management of "blinding" to medication effects.

Though review does not imply any necessary benefit from combining medication with psychological interventions, ensuring access to both seems warranted on pragmatic grounds. Clinically, it is not unusual for individuals who might benefit from psychological intervention to receive exclusively medical treatment; equally, patients receiving psychological treatment may be undermedicated or have received no psychiatric assessment when this might have been indicated. Better integrated treatment provision has a number of advantages. It would not only facilitate greater patient choice but greater efficacy is also likely, since it should ensure that patients are not treated over long periods with methods to which they are not responsive, or to which their response could be optimized.

The pattern of results for a structured therapy (and for pharmacotherapy) is quite consistent; a simple rule of thumb would be that, in around 50% of cases, symptoms will have remitted posttherapy, but that over 1 year of follow-up, around half of those who recovered will relapse. On this basis, only about one-fourth of patients treated using a brief therapy remain well. It is possible that the consistency of outcome reflects the fact that within any research sample, a proportion of patients will be responsive to almost any intervention. One strategy would be to screen out such individuals on the basis of their response to therapy, and to examine outcomes in those who appear to be treatment-resistant. Without this, there may be little to learn about the impact of specific techniques or, indeed, the differential impacts of medication and psychotherapy. However, this approach has the obvious drawback that defining a sample in terms of failure to respond to treatment does not necessarily guarantee its homogeneity: Treatment failure is under-

pinned by an admixture of biochemical and psychological factors, and without a hypothesized mechanism, we risk reifying what is only a description of outcomes into a unitary category. One approach to this problem is entirely pragmatic: broadly, a form of stepped care in which those who do not respond to one treatment are offered another, with the intent of maximizing outcomes. While this makes clinical sense, it is worth observing that developments in the field require a better understanding of the ways that different pathways to and presentations of depression contribute to outcome, something that can only be done in the context of hypotheses about the nature of the problems confronting depressed individuals.

Concerns about the limitations of brief intervention should not obscure its short-term benefit, but it is clear that treatment planning for individuals with depression should consider the need for maintenance therapy. In this respect, chronicity and age of onset appear to be more relevant factors to consider than severity. There is a risk that because the literature largely examines brief therapies delivered in single episodes of care, service provision will reflect this. A cascade or stepwise model may be more appropriate, with patients who fail to respond to brief (or even computerized) treatments offered alternative therapies, with the delivery of maintenance therapies seen as normative rather than unusual. The near-uniformity of treatment length in trials makes it very difficult to be clear about the duration of therapy. On this basis, there is little that can be said about the value of long-term therapy, though it is clear that very brief treatment regimens (of around 10 sessions) may not be adequate for more severely depressed individuals.

In summary, while treatments for depression are effective in the short term for at least a proportion of patients, longer term impacts are limited. Given the nature of the disorder, this is a creditable achievement, but it is clear that there needs to be a focus on the most pernicious aspect of life with this disorder—the tendency to relapse. Hand in hand with this, there may be a need to adopt a more complex framework for classifying depression. In this respect, greater consideration of developmental pathways, as well as personality variables, may be relevant factors for future research to consider.

Copyright © 2005 The Guilford Press. All rights reserved under International Copyright Convention. No part of this text may be reproduced, transmitted, downloaded, or stored in or introduced into any information storage or retrieval system, in any form or by any means, whether electronic or mechanical, now known or hereinafter invented, without the written permission of The Guilford Press.

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com