Assessing the Comparative Effectiveness of Antidepressant Therapies: A Prospective Clinical Practice Study

Gordon Parker, M.D., Ph.D., D.Sc, F.R.A.N.Z.C.P.; Kay Roy, B.A.; Kay Wilhelm, M.D., F.R.A.N.Z.C.P.; and Philip Mitchell, M.D., F.R.A.N.Z.P.

Background: Although efficacy studies suggest equal potency among antidepressant treatments, their effectiveness in clinical practice appears more variable, particularly in that the newer antidepressants may be less effective in either more severe depression or the melancholic subtype of depression. We pursue some factors that may impact the effectiveness of antidepressant treatments in a clinical sample.

Method: A sample of 182 patients with DSM-IV major depressive disorder was assessed at baseline and 12 months later to establish treatments provided, identify patients who had recovered from the index episode, and quantify likely treatment determinants. Four systems for distinguishing patients with melancholic and nonmelancholic depression were examined to assess for differential effects of the antidepressant strategies across those subtypes.

Results: Multimodal therapy (commonly, psychotherapy combined with an antidepressant drug) and patients' frequent attribution of recovery to spontaneous improvement made for difficulty in disentangling recovery determinants. After excluding a spontaneous improvement component, electroconvulsive therapy (ECT) and the irreversible monoamine oxidase inhibitors (MAOIs) appeared to be the most effective therapies across the sample, while the reversible inhibitor of monoamine oxidase-A (RIMA) appeared to be the least effective. The distinct gradient of suggested effectiveness of various strategies appeared to be contributed to principally by the varied effectiveness of alternate treatments across the melancholic subtype, whereby ECT, tricyclic antidepressants, and MAOIs were the most effective, and the selective serotonin reuptake inhibitors (SSRIs), RIMAs, and antipsychotic drugs were much less effective. For the nonmelancholic disorders, the effectiveness of SSRIs appeared to be comparable with that of older antidepressants.

Conclusion: Although most patients received a physical treatment, they commonly judged psychotherapy and spontaneous improvement to be influential in their recovery. Reasons for such attributions are worthy of clarifying studies. Despite patients' concerns about the side effects and stigma of ECT as well as the side effects associated with the older antidepressants, these therapies were rated as more helpful by patients—and were more strongly associated with recovery—than the newer antidepressant drugs. Such overall results are compatible with an earlier study undertaken by us involving an independent sample and retrospective data. The overall gradient is clarified by studying depressive subtypes, allowing an important conclusion. Although the newer and older antidepressant drugs may be of similar effectiveness in nonmelancholic depression, the newer agents appear comparatively inferior for the treatment of melancholia, findings that have clinical implications and perhaps inform us about the pathogenesis of melancholia.

(J Clin Psychiatry 2001;62:117–125)

Received June 25, 1999; accepted June 14, 2000. From the School of Psychiatry, The University of New South Wales, Randwick, Australia.

Supported by the National Health and Medical Research Council (Program Grant 993208).

Reprint requests to: Gordon Parker, M.D., School of Psychiatry, The University of New South Wales, Prince of Wales Hospital, High Street, Randwick, 2031, Australia.

Increasingly, there is a demand for and an acceptance of an evidence-based approach in psychiatry to assessing comparative benefits of various treatments. The evidence base for the efficacy of antidepressant drugs is substantive and seemingly clear-cut. For instance, a very recent review by the United States–based Agency for Health Care Policy and Research (AHCPR)[1] considered more than a thousand published trial reports and came to several important conclusions, particularly in regard to the newer antidepressants and major depression. First, more than 80 placebo-controlled studies found a response rate of 50% for the newer antidepressants (compared with 32% for placebo) for major depression. Second, the AHCPR report concluded that the newer antidepressants are as effective as the first and second generation tricyclic antidepressants (TCAs). Janicak et al.[2] earlier evaluated some 400 studies of the older and the more recent antidepressants (i.e., selective serotonin reuptake inhibitors [SSRIs], reversible monoamine oxidase inhibitors [RIMAs], monocyclic drugs, and tetracyclic drugs) and also concluded that the latter had comparable efficacy to their predecessors for major depression.

Such data build to the frequently heard view that all antidepressants are equally effective. In recent years, however, several reviews and studies[3–6] have suggested that the SSRIs may be less effective than the TCAs in treating melancholic depression, while another study[7] also challenged whether moclobemide was as effective as a TCA for melancholia. Clinicians also comment on such possible differences in effectiveness of many newer antidepressants.

A number of theoretical and practical reasons exist for potential dissonance between clinical trial data and "real world" effects. First, clinical trials that generate efficacy data are mostly restricted to patients with a rather diffuse and severity-based disorder (i.e., major depression). The trials most commonly involve outpatients without significant comorbidity (particularly personality disorder and drug and alcohol problems), and outcome is generally assessed dimensionally, rarely beyond 6 weeks. In clinical practice, nonspecific therapeutic factors (e.g., placebo and spontaneous remission) may have a major influence, disorder severity may differ considerably (more severe in hospitalized patients), patient motivation and treatment compliance will vary, comorbidity is common, and patients and therapists assess outcome in terms of recovery rather than percentage improvement. Wells[8] has recently drawn attention to such differences and to how efficacy data from clinical trials may differ from effectiveness estimates (the latter referring to outcome under conditions approximating usual care).

Diagnostic systems may also influence data interpretation. If the depressive subtypes have varying response patterns to differing antidepressant strategies, then diagnoses that measure severity (e.g., major depressive disorder) may obscure true differences across those subtypes. That interpretation is best illustrated in regard to psychotic (or delusional) depression. Two meta-analyses[9,10] have established that use of an antidepressant or an antipsychotic alone is effective in only a minority of patients, whereas a combination of the two and unilateral or bilateral electroconvulsive therapy (ECT) each have superior effectiveness rates. The National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program study[11] provides a second example. There, 250 outpatients with major depressive disorder were randomly assigned to a 16-week treatment involving 2 psychotherapies (cognitive-behavioral therapy [CBT] and interpersonal therapy [IPT]), a TCA (imipramine), or clinical management. Significant and similar improvement occurred across all treatment types, so that in the overall sample, no differential treatment effects were demonstrated. Superiority of 1 treatment (imipramine) was demonstrated, however, when analyses were limited to a subsample with more severe depression[12] and to a subsample in which a greater representation of any melancholic depressive subtype was suspected.

In a recent clinical panel study,[13] we asked depressed patients, who were under psychiatric treatment, in Australia and New Zealand to judge the effectiveness of previously received antidepressant treatments. Here, the overall sample gradient suggested high effectiveness for ECT and antipsychotic medication, moderate effectiveness for TCAs and irreversible monoamine oxidase inhibitors (MAOIs), and lesser effectiveness for a number of the newer antidepressants, including the SSRIs. Direct comparison of the SSRIs and TCAs indicated that both drug classes were similarly effective for nonmelancholic depression, but that the TCAs were distinctly more effective for melancholic depression, suggesting that any differential effectiveness may be distinctly contributed to by depressive subtype.

In this article, we prospectively study a sample for which treatment was based on individual clinician judgment, and we examine issues associated with outcome in the clinical setting that are generally not addressed or that are "controlled out" in formal efficacy studies. In a sense, we pursue an issue that confronts clinicians (i.e., Why did my patient improve or not improve with my management strategy?). Thus, we seek to make some estimate of the comparative effectiveness of various antidepressant strategies in a clinical setting where therapy (whether single or multiple, or involving pluralistic treatment modalities) as well as nonspecific improvement factors often confound interpretation.

Our second objective is to reexamine the issue of whether melancholic and nonmelancholic subtypes show differential responsiveness to treatments in the clinical setting. This raises the question of how melancholia may be most validly measured. In this article, we differentiate (nonpsychotic) melancholic and nonmelancholic depressive disorders using 4 diagnostic systems.

METHOD

Sample Recruitment and Baseline Assessment

We have previously described intake[14] assessment procedures, and thus now cover issues germane to this report. Sample members were grouped into several sets of depressed patients. The first set comprised 204 consecutive referrals to our tertiary Mood Disorders Unit (MDU), a sample weighted somewhat to those with severe and/or treatment-resistant conditions. To redress such influences, we also recruited 66 depressed patients treated by the authors as outpatients or as hospitalized catchment area patients. All were required to have a primary DSM-III-R major depressive disorder present for less than 2 years, with the duration criterion seeking to exclude persistent treatment resistance. MDU assessment generally resulted in recommendations to the referring psychiatrist or primary practitioner, while for the remaining subjects, the assessing psychiatrist generally recommended at least ini-

tial treatment. Thus, sample members were treated according to the clinical judgment of a responsible clinician, and the study sought to examine the effectiveness of such treatment over the following 12 months.

Baseline severity of depression was assessed by the clinician-rated Hamilton Rating Scale for Depression (HAM-D)[15] and by the self-reported Beck Depression Inventory (BDI).[16] A semistructured interview and self-report sheets assessed a wide range of symptoms,[14] and the clinician rated the patient on the formalized CORE[17] measure of observable psychomotor disturbance, which provides an estimate of the probability of melancholia.

DSM-IV diagnoses were generated, allowing 3 principal groups (i.e., major depressive disorder alone, major depressive disorder with melancholia, and major depressive disorder with psychotic features) to contrast any differential effect of treatment on melancholic and nonmelancholic depression. We also assigned those (nonpsychotic) patients with the preestablished CORE cutoff score of 8 or more to a CORE-defined melancholic group. Within our current sample, we developed a latent class analysis–derived algorithm of clinical features[18] for assigning (nonpsychotic) patients to a melancholic class (contributed to principally by a high CORE score, as well as by certain endogenous depression symptoms such as loss of interest, nonreaction, and nonvarying mood), also allowing us to assign patients to melancholic and nonmelancholic groups. Finally, at intake assessment, psychiatrists were given descriptors of clinical diagnoses for assigning patients to 4 categories,[17] again allowing a nonpsychotic melancholic group (the clinically labeled "endogenous" group) initially to be contrasted with a residual nonmelancholic group (combining the clinically labeled "neurotic" and "reactive" groups).

Assessment at the 12-Month Review

At the follow-up, the interviewing research assistant and the consultant psychiatrist used a structured format to determine details about treatments received over the interval and the outcome of the index episode. Progress over the 12 months was assessed principally by applying slightly modified Frank et al. criteria[19] that operationalize change point definitions. Thus, *partial remission* was defined as a spontaneous or treatment-induced state in which the patient has shown distinct improvement and no longer meets major depressive episode criteria, but has some symptoms; *relapse* as a change from either partial or full remission to full syndrome major depressive episode; *recurrence* as a return of symptoms following recovery, with symptoms of sufficient magnitude to qualify for a new major depressive episode; *full remission* as a spontaneous or treatment-induced state where patients become completely asymptomatic for less than 2 months; and *recovery* as a spontaneous or treatment-induced asymptomatic state lasting more than 2 months.

We established[20] that those change point definitions corresponded well with several assessments at the 12-month review (viz., a clinical global improvement score, no longer meeting criteria for DSM-IV–defined major depressive disorder, and a 50% reduction in BDI scores). For outcome of the index episode, we here combine the recovery and full remission categories as our "recovered" group, which means that the patient had had a complete disappearance of the major depressive disorder for 2 months or more following baseline assessment and treatment.

At follow-up, the psychiatrist was then required to determine (by questioning the patient as well as by clinical assessment) the most important factor in their improvement and, if necessary, the second and third most important factors—now termed "helpful interventions." Listed options included drug treatments, ECT, and a "spontaneous" option.

Determination of Improvement Factors

We sought to identify factors contributing to recovery, a process presenting a number of theoretical and practical difficulties when many variables are not controlled. Not surprisingly, patients receiving multiple drugs or multimodal treatments commonly had difficulty in nominating the rank order of "helpful interventions." Against expectation, the majority of patients (68%) judged that spontaneous improvement had occurred—to some degree or totally. When so reported, their judgment of this issue was respected unless the assessing clinician had strong clinical evidence to override a patient's judgment, despite the great majority of patients having been prescribed an antidepressant drug and/or ECT. Thus, the study essentially relied on the patients' estimates of "helpful interventions."

As a consequence of such issues, we adopted the following decision rules for data analyses:

1. Since most patients received more than one treatment and since disentangling any separate effect would be unsatisfactory for those receiving multitreatment regimens, each specific treatment was analyzed as a separate intervention.
2. For a treatment to be rated as effective, it must have been nominated as a helpful intervention and the patient must have recovered. Thus, even if an intervention or factor was judged as responsible for improvement, if the patient did not meet recovery status, the treatment was rated as ineffective.
3. If a patient had taken more than one drug of the same class (e.g., 2 SSRIs), with only one being judged as a "helpful intervention," we rate only that drug in terms of treatment effectiveness.
4. If a drug was withdrawn because of side effects, we rate it here as "not used" if it was ceased

Table 1. Helpfulness and Effectiveness of Specific Treatments (N = 182)[a]

Specific Treatment	Patients Receiving a Specific Treatment, %[b]	Helpful Intervention, %[c]	Treatment Effectiveness, %[d]	Attribution of Responsibility for Recovery[e]			
				This Treatment Alone, %	This Treatment Plus Another Treatment, %	Other Treatment Responsible, %	Spontaneous Remission Plus Treatment, %
Psychotherapy	78	70	47	6	26	9	54
SSRI	39	48	29	1	25	23	56
TCA	37	54	38	0	29	15	52
ECT	18	90	59	6	50	3	38
RIMA	12	27	18	0	14	18	64
Antipsychotic	11	38	35	0	19	19	38
MAOI	10	56	53	6	33	11	17

[a]Abbreviations: ECT = electroconvulsive therapy, MAOI = monoamine oxidase inhibitor, RIMA = reversible inhibitor of type A monoamine oxidase, SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant.
[b]Percentage of sample receiving that specific treatment, either alone or in conjunction with other therapies.
[c]Of those receiving each specific treatment, the rate of nominating it as a helpful intervention.
[d]Treatment effectiveness examines those rated as recovered/fully remitted who nominated the specific treatment as helpful.
[e]Attribution of responsibility for recovery expressed as a percentage of users of the specific treatment.

shortly after initiation and as "used" if it was taken for at least several weeks (this influences the denominator for those receiving each treatment).

5. Any treatment ceased because of judged clinical ineffectiveness is accepted here as "ineffective" (again influencing the denominator for each treatment group).

RESULTS

Of the 270 patients initially assessed at baseline, 67% completed the 12-month follow-up and make up our current sample. Those declining follow-up did not differ by mean age, sex, clinical diagnostic distribution, age at first depressive episode, lifetime duration of depression, or BDI and HAM-D depression severity measures. Patients who declined follow-up were less likely to be married or in a stable relationship and had had more (19.2 vs. 11.5, $t = 2.3$, $p < .05$) lifetime depressive episodes.

For the 182 patients assessed at follow-up, baseline data established a mean age of 44.4 years, a female preponderance of 62%, and a mean duration of 8.2 months for the current depressive episode. Using the change point indicators,[19] 57% met criteria for recovery and 5% for a full remission (i.e., in our analyses, 62% recovered from their episode). Over the 12 months, all patients received 1 or more active treatments. Specifically, sample members received a mean number of 2.4 therapies involving a mean of 1.6 medications (including lithium). Although 20% of the sample received lithium for mood stabilization or to augment their antidepressant, we did not include lithium in the data set due to the 2 contrasting rationales for prescription.

Frequency of Treatment Use

Table 1 (first column) reports the frequency of using the contrasting specific antidepressant treatments across the follow-up sample. Such data are informative only to the extent that they demonstrate treatments recommended either by the MDU consultants or by the clinicians responsible for ongoing management of the index episode. Although the data indicate that some form of psychotherapy was a very common treatment modality, a subsidiary analysis established that, for 87% of the psychotherapy patients, it was provided in conjunction with at least 1 drug treatment. For the purposes of this study, the term *psychotherapy* ranged across the formal psychotherapies to include counseling and the mere provision of advice, and so must be regarded here as a diffuse treatment category comprising both specific and nonspecific applications. Table 1 data also document the percentage of patients who received each treatment, including 4 antidepressant drug classes (SSRI, TCA, RIMA, and MAOI), and a subsidiary analysis established that 80% of the sample received 1 such antidepressant or more. Thus, the great majority of patients received both an antidepressant drug and psychotherapy.

Effectiveness of Specific Treatments

Table 1 also lists the percentage of patients receiving each treatment who nominated the treatment as a helpful intervention (second column) and quantifies the rate of recovery for patients who received each treatment and judged it to be a "helpful intervention" (third column). The remaining columns examine attributions from patients who recovered (i.e., for those receiving each treatment, the respective percentages of patients who rated it alone as effective, effective in combination with another "helpful intervention," or less effective than another nominated "helpful intervention" or who viewed the episode recovery as occurring spontaneously—entirely or to some degree).

Seventy-eight percent of patients received psychotherapy. Of those, 70% nominated it as a helpful intervention, and 47% not only judged it as a helpful intervention but also recovered. Six percent of the psychotherapy pa-

tients judged recovery to be due to psychotherapy alone, 9% judged another specific treatment as determining recovery, 26% judged psychotherapy plus another specific treatment as determining recovery, and 54% judged that recovery had a spontaneous component. Such analyses allow contrasting interpretations, with data suggesting at first glance that psychotherapy was quite effective. The "deconstructed" data suggest, however, that improvement while receiving psychotherapy may reflect spontaneous improvement or the impact of another specific therapy.

The data allow the treatments to be ranked in terms of helpfulness, ranging from highly helpful interventions (including ECT and psychotherapy) to moderately helpful interventions (the MAOIs, TCAs, and SSRIs) to somewhat helpful interventions (antipsychotics and the RIMA). Again, the degree to which helpful interventions were associated with episode recovery ranged considerably, being highest for ECT, MAOIs, and psychotherapy and lowest for SSRIs and the RIMA.

The rate at which remission was viewed as having a spontaneous component while patients were receiving a nominated helpful intervention ranged considerably, from 17% for an MAOI to 64% for the RIMA. Amalgamating 2 columns effectively excluded any spontaneous improvement component (and so assessed recovery due to 1 or more specific treatments alone), producing another distinct gradient, with ECT returning the highest rate (56%), followed by MAOIs (39%), psychotherapy (32%), TCAs (29%), SSRIs (26%), antipsychotics (19%), and the RIMA (14%).

Such data suggest, but do not prove, differential effectiveness of differing treatments across a sample with heterogeneous depressive conditions (i.e., psychotic, melancholic, and nonmelancholic). If this hypothesis is correct, it could reflect either the treatments' overall antidepressant potency or treatment specificity effects. The latter explanation is pursued via a focused comparison.

Differential Treatment Effectiveness in Melancholic and Nonmelancholic Depressive Subgroups

As noted in the introduction, we hypothesized that each of the examined treatments has variable effectiveness specificity across melancholic and nonmelancholic subtypes. Outcome is again judged in relation to episode recovery rates. Since there were few patients with psychotic depression (N = 18 at follow-up), we deleted them from our analyses and contrasted variably defined melancholic and nonmelancholic depressive subgroups.

Table 2 first reports overall effectiveness rates for the various treatments (differing somewhat from Table 1 data, since the patients with psychotic depression are now deleted). Nevertheless, the gradient across treatments is, apart from antipsychotic medication, exactly the same, with the highest effectiveness rate being for ECT and the lowest rate—not surprisingly, with the psychotic

Table 2. Percentages of Melancholic and Nonmelancholic Patients Rating Treatment Outcome as Effective According to Differing Definitions of Melancholia, in Response to Treatments Used Alone or in Conjunction With Other Treatments[a]

			Clinical				DSM-IV				Algorithm				CORE ≥ 8			
Treatment	N	Overall Effectiveness %	Melancholic (N = 60), %	Nonmelancholic (N = 105), %	OR	95% CI	Melancholic (N = 56), %	Nonmelancholic (N = 109), %	OR	95% CI	Melancholic (N = 63), %	Nonmelancholic (N = 93), %	OR	95% CI	Melancholic, (N = 37), %	Nonmelancholic, (N = 127), %	OR	95% CI
SSRI	70	29	26	30	0.83	0.27 to 2.55	20	33	0.50	0.16 to 1.59	30	27	1.17	0.39 to 3.54	23	30	0.71	0.17 to 2.89
TCA	59	34	53	26*	3.22	1.02 to 10.20	48	25	2.75	0.90 to 8.37	50	21*	3.83	1.17 to 12.51	43	29	1.79	0.52 to 6.19
MAOI	14	43	43	33	1.50	0.16 to 14.42	50	37	1.67	0.19 to 14.27	60	43	2.00	0.19 to 20.61	60	33	3.00	0.31 to 28.84
RIMA	22	18	33	12	3.50	0.37 to 33.31	25	17	1.67	0.13 to 22.00	22	17	1.43	0.16 to 12.70	33	16	2.67	0.18 to 39.63
ECT	23	56	67	50	2.00	0.29 to 13.74	50	67	0.50	0.09 to 2.84	69	33	4.50	0.57 to 35.52	67	45	2.40	0.44 to 12.98
Antipsychotic	12	17	11	33	0.25	0.01 to 5.98	14	20	0.67	0.03 to 14.03	14	0		NA[b]	40	0		NA[b]
Psychotherapy	133	46	45	47	0.90	0.42 to 1.92	42	48	0.79	0.37 to 1.66	43	51	0.72	0.35 to 1.49	41	48	0.76	0.30 to 1.92

[a]Abbreviations: CI = confidence interval, ECT = electroconvulsive therapy, MAOI = monoamine oxidase inhibitor, OR = odds ratio, RIMA = reversible inhibitor of type A monoamine oxidase, SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant. Because the number of patients with major depressive disorder with psychotic features was small (N = 18), they are excluded from this analysis.
[b]NA = Not appropriate to calculate odds ratio because of empty cell.
*$p < .05$ (chi-square analysis).

depressed patients deleted—now for antipsychotic medication.

Our focus, however, for Table 2 data is on the recovery rates (for each specific treatment) for patients assigned to melancholic and nonmelancholic categories. The use of 4 different diagnostic systems for subdividing melancholic and nonmelancholic depression allows for consistency to be examined across systems. In that regard, there are consistent trends (quantified by the odds ratios [ORs]) for patients assigned to the 4 melancholic categories to report a greater chance (than the contrasted nonmelancholic subjects) of recovery while receiving a TCA, an irreversible MAOI, or the RIMA. The only significant finding was in relation to the TCAs in which case, against 2 diagnostic system criteria, melancholic subjects were significantly more likely to have recovered—a finding confirmed by confidence interval estimates. This phenomenon was not evident for the SSRIs and was inconsistent in relation to ECT (failing to hold in relation to DSM-IV assignment).

The most effective treatments for melancholic depression (as variably defined) were ECT and the irreversible MAOIs and TCAs, with psychotherapy intermediate and with the RIMA, SSRIs, and antipsychotics rating relatively consistently as distinctly less effective across all diagnostic systems (see Table 2). For nonmelancholic depression, the range in effectiveness across treatments is less striking, but psychotherapy, ECT, and MAOIs appear the most effective, the SSRIs and TCAs intermediate, and the antipsychotics and the RIMA least effective.

The comparative analyses also have some potential to inform us about the utility of the 4 diagnostic systems. DSM-IV decision rules for melancholia rely very much on symptom data, and subjects assigned as melancholic or nonmelancholic by that system showed no significant difference in relation to any of the treatments considered. Use of the CORE score alone also failed to demonstrate any significant differences. Clinical diagnostic allocation and the algorithm, which both weight signs of psychomotor disturbance as well as endogenous symptoms to the definition of melancholic depression, did generate some significant differences for the TCAs, suggesting their greater effectiveness for melancholic depression. As subject numbers contribute to achieving or not achieving formal significance, inspection of odds ratios can be informative. Here we see a consistent trend for higher odds ratios in favor of a treatment being more effective for melancholia for TCAs (ORs = 1.8–3.8), MAOIs (ORs = 1.5–3.0), the RIMA (ORs = 1.4–3.5), and some inconsistency for ECT. By contrast, there is no clear differentiation for the SSRIs or for psychotherapy.

DISCUSSION

Efficacy studies have distinct advantages, particularly in being able to control a number of variables so that the placebo effect and other nonspecific effects as well as the efficacy of the active ingredient can be quantified. Such advantages should not be minimized. Nevertheless, there are obvious limitations to such controlled studies, in both their restrictions concerning eligible subjects (e.g., those with significant comorbidity problems are usually excluded) and the very artificiality of any trial. For example, patients or subjects will be aware of trial nuances (from briefings and review of consent forms), may guess whether they are receiving a particular active drug (or the placebo), and may wish to "please" the investigators for a range of reasons. Such factors, and many others, may limit the generalizability of the information to the clinical front. This does not, of itself, argue by default for uncontrolled studies of clinical samples, since characteristics of the sample, multiple uncontrolled variables, and other factors can lead to invalid conclusions in open studies.

While respecting that caveat, and conceding that our sample may or may not represent the world of clinical depression (and that relatively small numbers of patients received each designated specific treatment), we report a "real world" effectiveness study. Because of recruitment nuances, our sample may have been weighted to the more severe and treatment-resistant depressive disorders. Although a substantial minority did not accept follow-up review, we established that those patients did not differ distinctly from those accepting review (our current sample). Analyses did not evaluate dosage of medication, length of treatment, or other treatment details that might be expected to have some relevance. Again, although we had intended to rate treatment effectiveness by using both clinician and patient judgments, data collection relied almost entirely on patient reports if there was any discordance.

Self-reported information and patient-based attributions are clearly problematic, but even if not definitive, can be extremely salient. In another context, Jorm and colleagues[21] have studied the views of the general public and mental health professionals about helpful therapies for depression. While the professionals rated antidepressant drugs as highly helpful, the general public tended to rate self-help strategies more highly and viewed antidepressant medications as potentially addictive and both antidepressant drugs and ECT as harmful. Thus, attributions about helpful interventions range widely across professionals and patients.

We argue that the type of information assessed here (and in a previous clinical panel study[13] that we undertook) is of the caliber that frequently dictates clinicians' treatment decision making. Thus, it is desirable for clinicians to ask about previous treatments (and responses to them) when assessing any new patient, and presumably reports of any treatment having been previously effective or noneffective dictate the clinician's treatment decisions on such occasions. In essence, we built on such realities in our study design but then sought to deconstruct the impact

of differing major depressive types and several nonspecific treatment variables in evaluating differing specific treatments. Such an approach will not overcome the numerous confounds that limit interpretation but, as noted shortly, this concern can be addressed by examining for consistency and replication across independent studies and by reshaping controlled studies.

Table 3. Comparative Effectiveness Data (% Recovered) From Current Longitudinal Study (N = 182) and Earlier Retrospective Study[13] (N = 341) for Principal Treatments[a]

	Whole group		DSM-IV Melancholic		DSM-IV Nonmelancholic	
Treatment	Longitudinal Study	Retrospective Study	Longitudinal Study	Retrospective Study	Longitudinal Study	Retrospective Study
ECT	56%	56%	50%	57%	67%	48%
MAOI	43%	39%	50%	39%	37%	39%
TCA	34%	38%	48%	38%	25%	36%
SSRI	29%	31%	20%	22%	33%	36%
RIMA	18%	18%	25%	17%	17%	16%

[a]Abbreviations: ECT = electroconvulsive therapy, MAOI = monoamine oxidase inhibitor, RIMA = reversible inhibitor of monoamine oxidase-A, SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant. Psychotherapy's effectiveness was not examined in the retrospective study.

Our subjects were highly likely to receive both an antidepressant medication and some form of psychotherapy, while small percentages received ECT or an antipsychotic drug. Such information is, in itself, of limited importance since it may merely reflect the fashion of our consultants and other Australasian psychiatrists managing these patients. Turning to the patient-generated information examining the impact of 7 differing antidepressant interventions, ECT was rated as the most helpful treatment (despite our patients' commonly expressed concerns about side effects and stigma associated with the treatment). Psychotherapy was also rated highly. Finally, the older antidepressants (here, TCAs and MAOIs) were rated in terms of their helpfulness as equivalent to the newer antidepressants (here, SSRIs and the RIMA), despite their differential side effect profiles. Therapists are encouraged to consider cost-benefit issues in considering antidepressant choice and are frequently encouraged to weight the low side effect profile of the newer antidepressants. Our data suggest that patients may indirectly also make cost-benefit judgments and that ECT, TCAs, and MAOIs hold up strongly in this regard against some of the newer antidepressants.

Although the study findings could be idiosyncratic, they are supported by results from an earlier study[13] in which we surveyed an independent sample of depressed patients and reviewed the judged effectiveness of a range of antidepressant medications given to them for previous depressive episodes. That study sought retrospective data, whereas the present study was essentially longitudinal, with respective numbers of 341 and 182 depressed patients. In Table 3, we report results using the standardized DSM-IV diagnostic system included in both studies. For the whole sample, there is striking comparability in treatment effectiveness rates across the 2 studies and consistency in their rank ordering.

Here, for those who recovered and nominated a specific intervention as helpful, a clear gradient was evident across the treatments. ECT and psychotherapy again rated highly, while the older antidepressants (MAOIs, TCAs) rated slightly higher than the SSRIs and certainly higher than the RIMA. Factors other than differential drug effectiveness clearly may have contributed to such a gradient. The sample was heterogeneous, particularly in terms of depressive disorders, although our refined analyses excluded those with a psychotic depression. In terms of attributing reason for recovery, few subjects nominated only one specific treatment. Although all subjects received treatment, a substantial percentage judged that recovery occurred spontaneously (to a degree or totally), being most marked for those receiving the RIMA and least for those receiving ECT and antipsychotic medication. Since the current study is uncontrolled, this finding could reflect the latter treatments being given to those with disorders least likely to show a spontaneous remission, or it could reflect some valid judgment by patients about the true effectiveness of their treatment. The frequency of such an attribution should encourage exploratory studies since an explanation currently remains unclear. Possibilities include patients' judging that recovery is always spontaneous but that antidepressant treatments facilitate the process, viewing the treatment as having a placebo effect, or viewing depression as a time-limited disorder with specific treatments not playing much part in recovery. Further investigations of these findings would benefit clinicians in their day-to-day management of depressed patients.

Our second objective was to examine the narrower proposition that specific treatments may have varying effectiveness for melancholic and nonmelancholic depression. Although there were some diagnostic system–specific differences (e.g., ECT did not tend to be more effective with DSM-IV–defined melancholic depression, and use of the CORE cutoff score alone did not generate any significant difference across subtypes), some consistent trends were evident, although the relatively small cell numbers for a number of analyses risk misinterpretation.

The gradient in the overall effectiveness of differing antidepressant drug types across the combined melancholic and nonmelancholic subjects (from 43% for MAOIs down to 18% for the RIMA) was somewhat clarified by examining patterns across the subtypes. First, TCAs, MAOIs, and the RIMA were consistently superior for assigned melancholic (compared with nonmelan-

cholic) subjects. Second, and by contrast, the SSRIs appeared to have equal effectiveness rates for melancholic and nonmelancholic subjects. Third, within the assigned melancholic subjects, the TCAs and MAOIs were consistently more effective than the SSRIs and the RIMA, whereas such trends were not evident in nonmelancholic subjects. Confidence in these results is again encouraged by reference to the retrospective study.[13] In Table 3, we compare results for the 2 studies in relation to DSM-IV assignment to melancholic and nonmelancholic depression. As for the total samples, the compared subsamples show identical rank ordering of treatments—although here, percentage effectiveness rates do vary across the 2 studies. Ignoring ECT (which was suggested as the most effective treatment, independent of depressive subtype) and the RIMA (which returned low effectiveness rates across depressive subtypes), our current study suggests that the MAOIs and TCAs were twice as effective as the SSRIs (48%–50% vs. 20%) for melancholic depression whereas such drug class differentiation was not suggested for nonmelancholic depression (effectiveness rates ranging from 25% to 37%). Confidence in this result is aided by reference to the retrospective study (see Table 3) where, again, for melancholic depression, the TCAs and irreversible MAOIs returned effectiveness rates approximately twice as high as for the SSRIs but returned virtually identical effectiveness rates (of 36%–39%) for nonmelancholic depression.

We suggest, then, that the familiar claim that all antidepressants are equally effective should not be accepted at face value. Importantly, we find evidence that older antidepressants such as TCAs and MAOIs may be more effective in clinical practice than some of the newer ones (such as the SSRIs and RIMAs), but, most importantly, we suggest that any such differential effectiveness may be determined by the type (rather than the severity) of depression being treated. Our 2 studies suggest that any greater effectiveness of the older antidepressants is most evident in patients with melancholic depression whereas differential effectiveness may be less clear or absent in nonmelancholic depression. Since most efficacy studies are likely to involve less severely depressed patients (and have a low percentage of melancholic subjects), it is hardly surprising that they would fail to demonstrate differential effectiveness across the newer and older antidepressants. Thus, just as absence of proof is not necessarily proof of absence, we should not accept unchallenged the large synthetic reviews (e.g., the AHCPR report[1]) that conclude that the newer antidepressants are equally efficacious as first and second generation antidepressants, particularly if their contributing samples neither ensure representation of patients with a melancholic depression nor examine for differential effects in that subtype.

If the newer antidepressants such as the SSRIs and RIMAs are not as effective as the TCAs and MAOIs for treating melancholia, this finding has important clinical implications. Potentially, the differing actions of these antidepressants could hold the key to clarification of perturbed neurotransmitter mechanisms in melancholia, with clear research implications.

We conclude, then, that this sample examining clinical effectiveness provides data that allow certain hypotheses to be derived and shaped for refined analysis in controlled studies, particularly for examining differential effectiveness across melancholic and nonmelancholic depressive subtypes. Our results suggest also that the potential schism that exists between interpreting efficacy studies and effectiveness studies might be narrowed by study designs that move away from assessing depression on a dimensional basis (e.g., "major" and "minor" depression) to a more categorical basis, whether the latter contrast psychotic, melancholic, and nonmelancholic subtypes or adopt other operationalized categories. The difficulty with the latter approach is in ensuring valid subtyping. This issue can be addressed directly, or results of studies like this one can be used to assist in shaping disorder subtypes.

REFERENCES

1. Agency for Health Care Policy and Research. Treatment of Depression: Newer Pharmacotherapies [summary]. Report from University of Texas Health Science Center; San Antonio, Tex; 1999. Evidence Report No. 7
2. Janicak PG, Davis JM, Preskorn SH, et al. Principles and Practice of Psychopharmacotherapy: Management of an Acute Depressive Episode. Baltimore, Md: Williams & Wilkins; 1993:226–246
3. Controversies in the diagnoses and treatment of severe depression [ACADEMIC HIGHLIGHTS]. J Clin Psychiatry 1996;57:554–561
4. Roose SP, Glassman AH, Attia E, et al. Comparative effectiveness of selective serotonin reuptake inhibitors and tricyclics in the treatment of melancholia. Am J Psychiatry 1994;151:1735–1739
5. Danish University Antidepressant Group. Citalopram: clinical effect profile and comparison with clomipramine: a controlled multicenter study. Psychopharmacology (Berl) 1986;90:131–138
6. Danish University Antidepressant Group. Paroxetine: a selective serotonin reuptake inhibitor showing better tolerance, but weaker antidepressant effect than clomipramine in a controlled multicenter study. J Affect Disord 1990;18:289–299
7. Danish University Antidepressant Group. Moclobemide: a reversible MAO-A-inhibitor showing weaker antidepressant effect than clomipramine in a controlled multicenter study. J Affect Disord 1993;28:105–116
8. Wells KB. Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. Am J Psychiatry 1999;156:5–10
9. Spiker DG, Weiss JC, Dealy RS, et al. The pharmacological treatment of delusional depression. Am J Psychiatry 1985;142:430–435
10. Parker G, Roy K, Hadzi-Pavlovic D, et al. Psychotic (delusional) depression: a meta-analysis of physical treatments. J Affect Disord 1992;24: 17–24
11. Elkin I, Shea MT, Watkins JT, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. Arch Gen Psychiatry 1989;46:971–982
12. Elkin I, Gibbons RD, Shea MT, et al. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. J Consult Clin Psychol 1995; 63(suppl 5):841–847
13. Parker G, Mitchell P, Wilhelm K, et al. Are the newer antidepressant drugs as effective as established physical treatments? results from an Australasian clinical panel review. Aust N Z J Psychiatry 1999;33:874–881
14. Parker G, Hadzi-Pavlovic D, Roussos J, et al. Non-melancholic depression: the contribution of personality, anxiety and life events. Psychol Med 1998;28:1209–1219

15. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–296
16. Beck AT, Ward CH, Mendelson JE. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561–571
17. Parker G, Hadzi-Pavlovic D. Melancholia: A Disorder of Movement and Mood: A Phenomenological and Neurobiological Review. New York, NY: Cambridge University Press; 1996
18. Parker G, Wilhelm K, Mitchell P, et al. Sub-typing depression: testing algorithms and identification of a tiered model. J Nerv Ment Dis 1999;187: 610–617
19. Frank E, Prien RF, Jarrett RB, et al. Conceptualisation and rationale for consensus definitions of terms in major depressive episode: remission, recovery, relapse and recurrence. Arch Gen Psychiatry 1991;48:851–855
20. Parker G, Wilhelm K, Mitchell P, et al. Predictors of 1-year outcome in depression. Aust N Z J Psychiatry 2000;34:56–64
21. Jorm AF, Korten AE, Jacomb P, et al. Attitudes towards people with a mental disorder: a survey of the Australian public and health professionals. Aust N Z J Psychiatry 1999;33:77–83