Published in final edited form as:
Eur Neuropsychopharmacol. 2012 March ; 22(3): 183–199. doi:10.1016/j.euroneuro.2011.07.010.

Baseline depression severity as a predictor of single and combination antidepressant treatment outcome: Results from the CO-MED Trial

Edward S. Friedman[1], **Lori L. Davis**[2], **Sidney Zisook**[3], **Stephen R. Wisniewski**[4], **Madhukar H. Trivedi**[5], **Maurizio Fava**[6], **A. John Rush**[7], and **COMED Study Team**

[1]Department of Psychiatry, University of Pittsburgh School of Medicine, PA, USA

[2]Research and Development Service, VA Medical Center, Tuscaloosa, AL and Department of Psychiatry, University of Alabama School of Medicine, Birmingham, AL, USA

[3]Department of Psychiatry, University of California, San Diego and VA San Diego Health Care Systems, CA, USA

[4]Department of Epidemiology, University of Pittsburgh, PA, USA

[5]Department of Psychiatry, The University of Texas Southwestern Medical Center, Dallas, Texas, USA

[6]Depression Clinical and Research Program, Massachusetts General Hospital, Massachusetts, USA

[7]Office of Clinical Sciences, Duke-NUS, Singapore

Corresponding Author: Edward S. Friedman MD, 3811 O'Hara Street, Pittsburgh PA, 15213, Telephone: 412 246-5290, Fax: 412 246-5750, friedmane@upmc.edu.

STATEMENT OF INTEREST

Edward S. Friedman M.D. has received Grant/research support from Aspect Medical Systems, AstraZeneca, Bristol-Myers Squibb, Cyberonics, Indevus, Medtronics, Northstar, Novartis, Pfizer, Sanofi-Aventis, Wyeth-Ayerst, Repligen; and has served on the speakers bureaus or Advisory Boards for AstraZeneca, Lilly, GlaxoSmithKline, Pfizer, Wyeth-Ayerst, Bristol-Myers Squibb. Royalties: Springer.

Lori L. Davis M.D.Research support (no personal income received from these grants): Abbott Laboratories, Inc.; AstraZeneca; Bristol-Myers Squibb Company (BMS); Eisai Pharmaceuticals; Janssen; VA, NIMH, Shire; AstraZeneca; Southwestern Oncology Group, Department of Defense (DoD). Advisory/Consulting: Cyberonics; Abbott; Shire; Constella; Eli Lilly Speaking: Abbott Laboratories; Cyberonics; Sanofi-Aventis; AstraZeneca Equity Holdings (exclude mutual funds/blinded trusts): Pfizer (until 2009) Royalty/patent, other income: None

Sidney Zisook, M.D. has received grant support from National Institute of Mental Health, American Foundation for Suicide Prevention, the Department of Veterans Affairs and PamLab.

Stephen R. Wisniewski, PhD. Consulting: Cyberonic Inc. (2005–2009), ImaRx Therapeutics, Inc. (2006), Bristol-Myers Squibb Company (2007–08), Organon (2007), Case-Western University (2007), Singapore Clinical Research Institute (2009), Dey Pharmaceuticals (2010), Venebio (2010), Dey (2010).

John Rush, M.D., has received consultant fees from Advanced Neuromodulation Systems, AstraZeneca, Best Practice Project Management, Bristol-Myers Squibb/Otsuka, Cyberonics, Forest Pharmaceuticals, Gerson Lehrman Group, GlaxoSmithKline, Jazz Pharmaceuticals, Magellan Health Services, Merck & Company, Neuronetics, Novartis Pharmaceuticals, Ono Pharmaceuticals, Organon, Otsuka Pharmaceuticals, Pamlab, Pfizer, Transcept Pharmaceuticals, Urban Institute and Wyeth-Ayerst; speaking fees from Cyberonics Inc., Forest Laboratories, GlaxoSmithKline and Otsuka; royalties from Guilford Publications, and Healthcare Technology Systems, and research support from National Institute of Mental Health and the Stanley Medical Research Institute. He has owned shares of stock in Pfizer.

Abstract

The objective of this manuscript is to report associations between baseline depressive severity and (1) baseline sociodemographic and clinical characteristics, (2) treatment outcomes, and (3) differential outcomes for three treatment groups. Six hundred and sixty-five outpatients with nonpsychotic, major depressive disorder were prospectively randomized to treatment with either a selective serotonin reuptake inhibitor (SSRI) monotherapy (escitalopram plus placebo) or one of two antidepressant medication combinations (bupropion-sustained release plus escitalopram, or venlafaxine-extended release plus mirtazapine). For purposes of these analyses, participants were divided into four groups based on baseline severity by the 16-item Quick Inventory of Depressive Symptomatology - Self-Report (QIDS-SR_{16}) total score: mild (0–10) [N=81], moderate (11–15) [N=238], severe (16–20) [N=260] and very severe (21–27) [N=67]. Treatment outcomes at 12 and 28 weeks were compared among the four severity groups. A history of childhood neglect and/or abuse was strongly associated with the severity of adult depression (1/2 of participants in the very severy group versus 1/5–1/4 of those in the mild group reported abuse and/or neglect). The degree of suicidality (e.g., 15/.4% of the very severe group ever attempted suicide versus none in the mild group), the number of suicide attempts (e.g., mean of .41 +/− 1.99 suicide attempts in the severe group versus o.o +/−0.0 in the mild group) and severity of suicidality (e.g., 9.2% of participants in very severe group had a plan or made a gesture versus 5.6% in moderate group and none in the mild group) were increased in more severe groups. Participants with a greater baseline depressive severity reported significantly more psychiatric comorbitities (e..g. [at $p < 0.05$] increased rates of agoraphobia, bulimia, generalized anxiety, hypocondriasis, panic disorder, post-traumatic stress disorder, social phobia and somatoform disorder, with 23.9 % of participants in the very severe group having reported four or more psychiatric disorders versus 1.2% of the mild group). Combination medication treatments were no more effective in treating severe depressions than was SSRI monotherapy. Remission (61.7% of participants in the mild group achieved remission versus 28.4% in the very severe group) is more difficult to achieve in more severe groups than is response (48.8% of participants in the mild group achieved response versus 58.2% in the very severe group) ($p < 0.03$) . These data may help us to understand the impact of baseline features on antidepressant medication effectiveness and to inform the personalization of depression treatment across the spectrum of depressive severity.

Keywords

Depression; abuse; suicide; combination treatment severity; response; remission

INTRODUCTION

Pretreatment of depressive symptom severity is one of the most robust baseline predictors of antidepressant medication (ADM) treatment outcome. More severely depressed patients at baseline have less favorable outcomes (Van et al., 2008) and lower probabilities of achieving remission (Vallejo et al., 1991; Hollon et al., 1992; Tedlow et al. 1998; Joffe et al., 1999; Brown et al., 2000). The Sequence Treatment Alternatives to Relieve Depression (STAR*D) study (Trivedi et al., 2006), found that greater baseline depressive severity, greater psychiatric and medical comorbidity, and less social support were associated with lower remission rates to treatment with the antidepressant citalopram.

The large, multi-site, NIMH Treatment of Depression Collaborative Research Program trial (Elkin et al., 1995) found that initial severity of depression predicted differential treatment effects. Imipramine and clinical management were extremely effective for more severely ill patients compared to two types of psychotherapy. A meta-analysis (Khan et al., 2002) found that for patients treated with ADM, a higher initial depressive severity by the 17-item

Hamilton Rating Scale for Depression ($HRSD_{17}$) (Hamilton, 1960) was associated with a significantly greater magnitude of symptom reduction; while for patients treated with placebo, a higher initial severity was associated with a smaller reduction in symptoms. Furthermore, early discontinuation was more frequent among participants with high initial depressive severity. A more recent meta-analysis (Fournier et al., 2010) found that the magnitude of benefit from ADM treatment compared with placebo increases with the severity of initial depressive symptoms and, on average, may be minimal or nonexistent in patients with mild-to-moderate depression.

This report uses data from the Combining Medications to Enhance Depression Outcomes (CO-MED) study (Rush et al., in press), to evaluate the relationship between baseline depressive severity and ADM treatment outcome. Specifically, we addressed the following questions:

Aims of the Study

1. How are sociodemographic and clinical features related to baseline depressive severity?
2. What is the relationship between baseline depressive severity and treatment outcome(s)?
3. Is baseline depressive severity associated with different outcomes to single or combination antidepressant treatment?

METHODS

Study Overview

Methodological details of the study are available elsewhere (Rush et al., in press). The following is a brief overview.

CO-MED was a 7-month single-blind, randomized trial that compared the effectiveness of each of two ADM combinations (bupropion-sustained release [BUP-SR] plus escitalopram [ESCIT], or venlafaxine-extended release [VEN-XR] plus mirtazapine[MIRT]) against that of a selective serotonin reuptake inhibitor (SSRI) escitalopram (ESCIT) plus placebo (PBO) (1:1:1 ratio) at 12 and 28 weeks of first-step acute-phase MDD treatment.

Site Selection

Clinical sites were chosen to ensure (a) adequate patient flow, (b) committed administrative support, (c) adequate minority representation, and (d) adequate representation of both primary (n=6) and psychiatric care (n=9) sites.

Recruitment

Potential outpatient participants were screened at each clinical site using each site's standard procedure (variable across sites). Commercial advertising for the CO-MED trial was not used as a method of recruitment.

Participants

Broad inclusion and minimal exclusion criteria were used to ensure a reasonably representative participant sample. Outpatient enrollees, 18–75 years of age, met DSM-IV TR (American Psychiatric Association, 2000) criteria for either recurrent (≥1 prior major depressive episode [MDE]) or chronic MDD (current MDE for ≥2 years) based on a clinical interview and confirmed using a DSM-IV MDD symptom checklist completed by the

Clinical Research Coordinator (CRC). The index episode had to be ≥6 months in duration and the baseline $HRSD_{17}$ score had to exceed 15. See www.co-med.org for a complete listing of inclusion/exclusion criteria.

The study protocol was developed according to the principles of the Declaration of Helsinki. The study protocol and all consent and study procedures were approved by the Institutional Review Boards at the National Coordinating Center (The University of Texas Southwestern Medical Center at Dallas), the University of Pittsburgh Data Coordinating Center, each participating Regional Center, and all relevant clinical sites.

Baseline Data

Sociodemographic and illness features were gathered at baseline. The self-report Psychiatric Diagnostic Screening Questionnaire (PDSQ) (Zimmerman and Mattia, 2001a; 2001b) established the presence of current Axis I disorders with 90% specificity (Rush et al., 2005). The Concise Health Risk Tracking – Self-Report scale (Trivedi et al., submitted) established the presence of suicidal ideation, the Altman Self-Rating Mania Scale (Altman et al., 1997) established the presence of manic symptoms, and the Cognitive and Physical Functioning Questionnaire (Fava et al., 2009) measured functioning. The Self-administered Comorbidity Questionnaire (SCQ) (Sangha et al., 2003) established the presence, severity, and functional impact of a range of common general medical comorbidities.

Antidepressant Treatment

The primary analysis was conducted after 12 weeks of treatment. Secondary analyses were conducted at week 28. Treatment visits were planned at baseline and weeks 1, 2, 4, 6, 8, 10, 12, 16, 20, 24, and 28. Measurement-based care provided personalized and vigorous dosing (Trivedi et al., 2006; 2007; Trivedi and Daly, 2007), with dosage adjustments based on the 16-item Quick Inventory of Depressive Symptomatology – Clinician-rated (QIDS-C_{16}) (Rush et al., 2003; 2006; Trivedi et al., 2004) which was extracted from the 30-item Inventory of Depressive Symptomatology (IDS-C_{30}) (Rush et al., 1996), and the Frequency, Intensity and Burden of Side Effects Rating (FIBSER) (Wisniewski et al., 2006) obtained at each treatment visit, and was guided by the CO-MED Operations Manual (available at www.co-med.org).

Treatment was randomly assigned, stratified by clinical site using a Web-based randomization system (Wisniewski et al., 2004), with random block sizes of three and six. Dosing schedules were based on prior reports (Fava, 2001; Papakostas et al., 2005; McGrath et al., 2006; Leuchter et al., 2008). Doses were increased only in the context of acceptable side effects. Participants could exit the study if unacceptable or intolerable side effects occurred that could not be resolved with dose reduction or medication treatment of side effects.

ESCIT plus PBO—ESCIT began at one tablet (10 mg)/d; to be increased to two tablets (20 mg)/d at 4 weeks if the QIDS-C_{16} was >5 (side effects allowing). Pill PBO was started at week 2, with the option to increase to two pills at week 4 if the QIDS-C_{16} was >5 (side effects allowing).

BUP-SR plus ESCIT—BUP-SR (150 mg/d) was started at baseline and increased to 300 mg/d at week 1. ESCIT began at 10 mg/d at Week 2. At week 4, BUP-SR was raised to 400 mg/d and/or ESCIT was raised to 20 mg/d if the QIDS-C_{16} was >5 (side effects and tolerability allowing). At week 6 and beyond, doses were increased up to a maximum of BUP-SR 400 mg/d (200 mg/d b.i.d.) and ESCIT 20 mg/d if the QIDS-C_{16} was >5 (side effects permitting).

VEN-XR plus MIRT—VEN-XR began at 37.5 mg/d for three days and then was raised to 75 mg/d. At week 1, VEN-XR was raised to 150 mg/d. At week 2 (if QIDS-C_{16} >5), MIRT was added (30 mg/d). At week 4 (if QIDS-C_{16} >5), VEN-XR could be raised to 225 mg/d and/or MIRT was increased to 30 mg/d. At week 6 (if QIDS-C_{16} >5), MIRT could be raised to 45 mg/d (maximum dose). At week 8 (if QIDS-C_{16} >5), VEN-XR could be raised to 300 mg/day (maximum dose).

Medication Blinding

One medication in each treatment group was open label (both participant and study personnel unblinded), while one medication was blinded (participant only) throughout the 7-month study. In the ESCIT+PBO group, the PBO was blinded; in the BUP-SR+ESCIT group, ESCIT was blinded; and in the VEN-XR+MIRT group, MIRT was blinded. The CRCs and physicians were not blinded to the treatments in order to maximize safety and allow physicians to make informed flexible dosing decisions

Concurrent Treatments

Only protocol antidepressant medications were allowed. Other treatments with possible antidepressant effects were proscribed, as were depression-targeted, empirically-validated psychotherapies for depression. Other psychotherapies (e.g., supportive, couples, occupational therapy) were allowed, as were medications for any general medical comorbidity. Based on clinician judgment, medications to treat antidepressant medication side effects were allowed in order to mimic practice and enhance retention.

Research Outcomes

Outcome assessments were collected at baseline and all treatment visits. The primary outcome, symptom remission, was based on the 16-item Quick Inventory of Depressive Symptomatology – Self-Report (QIDS-SR_{16}) (Rush et al., 2003; 2006; Trivedi et al., 2004). Remission was ascribed based on the last two consecutive measurements obtained during the 12-week acute trial to ensure that a single “good week” was not falsely signaling remission. At least one of these ratings had to be <8, while the other had to be <6. If participants exited before 12 weeks, their last two consecutive QIDS-SR_{16} scores were used to ascribe remission. Those who exited before having two post-baseline measures were considered not remitted.

Participants could exit the study if they had received a maximally tolerated dose(s) for ≥4 weeks by week 8 without receiving at least a 30% reduction in baseline QIDS-C_{16}. They could enter continuation treatment (weeks 12–28) if they had received an acceptable benefit (defined as a QIDS-C_{16} ≤9 by week 12) or if they reached a QIDS-C_{16} of 10–13 with clinician and participant judging the benefit to be substantial enough to indicate a treatment continuation. Thus, virtually all participants entering the continuation phase had at least a 40% reduction in baseline QIDS-C_{16}. When participants exited the study at any time, a Study Exit Form was completed. The CRC attempted to contact all participants who did not come for a final exit visit.

Secondary outcomes included attrition, response (>50% reduction in QIDS-SR_{16} from baseline), side-effect burden as measured by the FIBSER, and specific side effects as measured by the Systematic Assessment for Treatment Emergent Events–Systematic Inquiry (SAFTEE-SI) (Levine et al., 1986; Levine and Schooler, 1992), change in anxiety as measured by the anxiety subscale of the IDS-C_{30} (Rush et al., 1996; 2000; Trivedi et al., 2004), function as measured by the Work Productivity and Activity Impairment scale (Reilly et al., 1993) and the Work and Social Adjustment Scale (WSAS) (Mundt et al., 2002), and

quality of life as measured by the Quality of Life Inventory (QOLI) (Frisch, 1994; Frisch et al., 2005).

Baseline depressive severity

Baseline depressive severity was defined by baseline scores on the QIDS-SR_{16}. The sample was divided into four severity categories by QIDS-SR_{16} total score: mild (0–10), moderate (11–15), severe (16–20), and very severe (21–27).

Statistical Analyses

Descriptive statistics, including measures of central tendency and dispersion, were computed for continuous data. Frequency distributions were estimated for categorical data. The appropriate parametric (e.g., t-test) or nonparametric test (e.g., chi-square, Wilcoxon tests) were used to assure a balanced distribution of the sociodemographic, psychiatric, and medical characteristics among those with mild, moderate, severe, or very severe depression scores at baseline.

At 12 and 28 weeks, regression models were used to compare unadjusted and adjusted outcomes among participants with mild, moderate, severe, or very severe depression scores at baseline. The type of regression models varied by outcome and included linear regression, logistic regression, ordinal logistic regression and negative binomial regression models. Potential confounders were identified using a stepwise logistic regression model with an indicator of mild, moderate, severe, or very severe depression scores at baseline as the outcome and all other baseline characteristics as independent variables. Those variables that remained in the final stepwise model were considered as potential confounders in the adjusted models. The moderating effect of baseline depression severity scores on treatment was evaluated on two outcomes, severity of depression (QIDS-SR_{16}) and side effect burden (FIBSER Burden), at 12 and 28 weeks. For severity of depression, a linear regression model was fit, and for side effect burden an ordinal logistic regression model was fit. Both models included main effects for treatment and baseline severity score, as well as the two-way interaction between treatment and baseline mild, moderate, severe, or very severe depression scores. All analyses are considered to be exploratory in nature and a type I error or p-value $<.05$ was used as a threshold to identify statistical significance. When a statistically significant effect of baseline depression severity was identified, pairwise comparisons were conducted with a Bonferroni correction for multiple comparisons. A number of outcomes were examined and no adjustments were made for testing multiple outcomes, so results should be interpreted accordingly.

RESULTS

Of 835 participants approached for the study, 734 (87.9%) provided written informed consent for screening. Of those screened, 665 (90.6%) were eligible and were randomly assigned to one of the three treatment groups. Of these, 81 (12.5%) had a mild level of baseline depressive severity by the QIDS-SR_{16}, 238 (36.9%) had moderate baseline severity, 260 (40.5%) had severe baseline severity, and 67 (10.4%) had very severe baseline severity.

Table 1 compares these four groups. Participants with greater baseline depressive severity were more likely to be younger, female and have a lower monthly household income, and were less likely to be employed.

Participants from the group with greater baseline depressive severity were more likely to have had their first depressive episode before age 18, to have attempted suicide, or have a greater lifetime severity of suicidality. The number of suicide attempts reported increased with increasing baseline depressive severity groups. Greater baseline depressive severity

was associated with a greater likelihood of neglect, emotional abuse, physical abuse, or sexual abuse before age 18. The age of physical abuse was generally younger as baseline severity increased (Table 2).

Participants with a greater baseline depressive severity reported significantly more agoraphobia, bulimia, generalized anxiety disorder, hypochondriasis, panic disorder, post-traumatic stress disorder, and social phobia (all approximately 5–20 times greater than in the mild group), and more psychiatric disorders in general . Participants with greater baseline depressive severity did not report significantly more alcohol or substance abuse ($p > 0.05$) (Table 3).

Participants with greater baseline depressive severity presented with higher rates of lethargic depression and sleep disturbance by the IDS-C_{30}; higher rates of anxious features, atypical features, and melancholic features; and much higher rates of suicidal thoughts and plans. Baseline depressive severity, as seen by QIDS-SR_{16} groups, was consistently reflected in $HRSD_{17}$, IDS-C_{30}, and QIDS-C_{16} scores. Participants with greater depressive severity had poorer quality of life (QOLI) and greater functional impairment (WSAS) (Table 4).

In general, the treatment features over 12 weeks were similar among the four groups. A few differences were identified, such as the very severely depressed group receiving significantly higher doses of venlafaxine than the mild group. Also, participants in the moderate and severe groups received significantly higher last mirtazapine doses vs. those participants in the mild group. At 28 weeks, there were no significant differences between severity groups regarding treatment features.

Regarding week 12 outcome measures, about one quarter of participants exited the study before 12 weeks. Participants in the more severe baseline depression groups were significantly more likely to reach response ($p=0.0285$), with those in the moderate, severe and very severe groups 1.370, 1.874, and 4.236 times more likely to achieve response, respectively, than those in the mild group. Adjusted post-hoc tests indicated significant pair-wise differences in “Percent QIDS-SR_{16} change” (after correcting for multiple comparison) in the severe vs. mild, very severe vs. moderate, and very severe vs. severe groups ($p < .0001$). “Percent QIDS-SR_{16} change” varied significantly between the severe vs. mild, very severe vs. moderate and very severe vs. severe groups. Additionally, greater baseline symptom severity was associated with poorer quality of life as an outcome ($p=0.0174$) (Table 5). After adjustment for potential confounders, there were no significant differences between severity groups on side effect measures (FIBSER), the number of psychiatric and non-psychiatric serious adverse events, or in rate of remission.

Regarding week 28 outcome measures, a little more than one-third of participants exited the study in the continuation phase between weeks 12 and 28. Participants with a very severe baseline depression were significantly more likely to reach response than those in the moderate and severe groups, with those in the moderate, severe and very severe groups 1.210, 1.353, and 5.548 times more likely to achieve response, respectively, than those in the mild group. There were significant pair-wise differences ($p=0.0003$) in the “Percent QIDS-SR_{16} change” (after correcting for multiple comparison) in the very severe vs. moderate and the very severe vs. severe groups, with a greater reduction seen in the more severe groups. Additionally, as baseline severity increased, participant quality of life decreased ($p=0.0234$) (Table 6). After adjustment for potential confounders, there were no significant differences between severity groups and outcomes (side effects, the number of psychiatric and non-psychiatric serious adverse events, or in rate of remission).

The moderating effect of treatment across the depression severity groups was examined at week 12 and there was no differential effect of treatment across the depression severity

groups with respect to early termination, response, remission, last FIBSER Burden, last QIDS-SR_{16} or percent of QIDS-SR_{16} reduction (Table 7).

DISCUSSION

Our results indicate several important findings: 1) childhood neglect and abuse is strongly associated with the severity of adult depression, 2) the degree of suicidality, the number of suicide attempts and severity of suicidality are increased in more severe groups, 3) participants in more severe baseline depression groups had significantly greater medical and psychiatric comorbidity, 4) combination medication treatments are no more effective in treating severe depressions than is SSRI monotherapy, and 5) remission is more difficult to achieve in more severe groups than is response.

There was a strong association between a history of childhood emotional, physical, and/or sexual abuse and the baseline severity of adult depression. More severely depressed participants were significantly more likely than those less severely depressed to have experienced neglect or emotional, physical or sexual abuse before age 18, and the age of reported physical abuse was generally lower as baseline severity increased. Previous studies have shown that individuals with a history of childhood abuse have a worse antidepressant response compared to those without a history of childhood abuse (Nemeroff et al. 2003), and a recent, comprehensive, systematic review reported much increased odds of adult depression in those who had experienced childhood sexual abuse (Chen et al. 2010).

Our finding that the degree of suicidality, the number of suicide attempts and severity of suicidality are proportionally greater as severity of depression increases has important clinical implications. The association of depression with suicide risk is a clinical mainstay (Robins, 1986) and the increase in suicide risk with increased depressive severity has been demonstrated in a large longitudinal cohort study (Bradvik et al. 2008). However, the current study is the first prospective clinical trial to make this observation.

In this effectiveness population, which was a highly chronic and/or recurrently depressed sample, baseline severity did not predict a differential response to monotherapy or combination treatment. Similarly, in their meta-analysis of the relationship between initial depressive severity and efficacy in FDA antidepressant trials, Kirsch et al. (2008) concluded that drug-placebo differences in antidepressant efficacy increase as a function of baseline severity, but that this difference is attributable to decreased placebo responsiveness among very severe individuals (i.e., there was little differential medication effect moderated by baseline severity). To personalize this result for the individual patient, this data suggests that despite initial baseline depression severity, treatment should be the simplest and most tolerable (and affordable) for the patient.

Finally, our results indicate that sustained remission is more difficult to achieve in more severe groups than is response. This confirms the findings of other studies (Vallejo et al., 1991; Khan et al., 2002) which also found that greater baseline severity was associated with greater symptom reduction with ADM treatment, as well as, confirming previous findings (Hollon et al., 1992; Tedlow et al., 1998; Joffe et al., 1999; Brown et al., 2000; Trivedi et al., 2006) that the probability of remission decreases as baseline depressive severity increases.

One of this study's strengths was the CO-MED design decision to use very broad inclusion criteria, which yielded a real-world population of chronic and recurrently depressed outpatients across the severity spectrum. Additionally, this population had significant medical and psychiatric comorbidity. Although the inclusion criteria mandated a baseline $HRSD_{17} \geq 16$ at entry, the mean $HRSD_{17}$ in the mild severity group was 19±3.1 and was 28±4.8 in the very severe group, so most of the participants in this study would have been

eligible to participate in most efficacy studies on the basis of depression severity. The broad range of depressive severity in this study's participants enabled us to examine response and remission characteristics after both 12 and 28 weeks of ADM treatment between categorically-defined levels of baseline depressive severity.

There appear to be correlates of illness severity and baseline features (e.g., medical and psychiatric comorbidity, lethargic, anxious, melancholic and/or atypical features, suicidal thoughts and/or plans, impaired social function), but, whether these correlates have any implication for cause cannot be addressed in this study. Baseline severity did not (in this highly chronic and/or recurrently depressed sample) predict a differential response to monotherapy or combination treatment. This study demonstrates that response clearly is easier to achieve with greater baseline severity. As for remission, a single assessment (the usual in the literature) may inaccurately relate to baseline severity. Utilizing the CO-MED study's more restrictive definition of remission (at least two consecutive weeks with QIDS-SR_{16} = 6/8) there was no difference between severity groups regarding remission. Because the mild group only achieved remission using the single assessment definition of remission suggests that the milder group may have had a "wobbly", inconsistent, remission (i.e., for chronically and/or highly recurrently ill individuals with mild/moderate levels of depressive severity, modest fluctuations of a few points on the QIDS-SR_{16} may effect whether they achieve or remain in remission).

This study had several limitations. First, the rating scales used to measure depressive severity may not have accurately done so. For example, compared to the QIDS-SR_{16}, the $HRSD_{17}$ includes additional dimensionalities which may not be helpful in differentiating response to ADM (Rush et al., 2006). The $HRSD_{17}$ rater-derived score of >16, used to determine study eligibility, may be artificially deflated as raters compare the participant to other participants with depression. Conversely, the self-rated QIDS-SR_{16} score may be artificially inflated because participants subjectively rate their depression based upon their own experience (Dunlop et al., 2010). Additionally, the degree-of-agreement between patient- and clinician-rated scales of depressive severity varies widely and are particularly poor prior to the initiation of treatment (Dunlop et al., 2010). Second, due to the small sample sizes resulting from the division of the study population into four severity groups, our analysis of the combination treatments by baseline severity may lack the statistical power needed to differentiate outcomes, leading to a Type I error. Third, the results found with the medication combinations we used may not be generalizable to other possible medication combinations, such as combination treatments at higher dosages, or the combination of an antidepressant and a second-generation antipsychotic medication. Fourth, we must always consider the possibility that nonspecific treatment factors and site differences influenced the outcome of this study.

In summary, these results indicate: 1) childhood neglect and abuse is strongly associated with the severity of adult depression, 2) the degree of suicidality, the number of suicide attempts and severity of suicidality are increased in more severely depressed groups, 3) participants with a greater baseline depressive severity reported significantly more medical and psychiatric comorbidities than those in the mild group, 4) combination medication treatments are no more effective in treating severe depressions than is SSRI monotherapy, and 5) remission is more difficult to achieve in more severely depressed groups than is response. The results of this study may help us to understand the impact of baseline features on antidepressant medication effectiveness and to inform the personalization of depression treatment across the spectrum of depressive severity.

Acknowledgments

The authors are entirely responsible for the scientific content of this paper. This study was supported and funded by NIMH grant MH-9008. We would also like to acknowledge the editorial support of Jon Kilner, MS, MA (Pittsburgh, PA). We acknowledge the administrative support of the Research and Development Services at the participating VA Medical Centers.

REFERENCES

Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. Biological Psychiatry. 1997; 42(10):948–955. [PubMed: 9359982]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington DC: American Psychiatric Press; 2000. text revision

Bradvik L, Mattison C, Bogren M, Nettelbladt P. Long-term suicide risk of depression in the Lundby cohort 1947–1997- severity and gender. Acta Psychiatrica Scandinavica 2008. 2008; 117:185–191.

Brown C, Schulberg HC, Prigerson HG. Factors associated with symptomatic improvement and recovery from major depression in primary care patients. General Hospital Psychiatry. 2000; 22:242–250. [PubMed: 10936631]

Chen LP, Murad MH, Paras ML, Colbenson KM, et al. Sexual abuse and lifetime diagnosis of psychiatric disorders: systematic review and meta-analysis. Mayo Clinic Proceedings. 2010; 85(7): 618–629. Epub 2010 May 10. [PubMed: 20458101]

Dunlop BW, Li T, Kornstein SG, Friedman ES, et al. Correlation between patient and clinician assessments of depression severity in the PREVENT study. Psychiatry Research. 2010; 177:177–183. [PubMed: 20304503]

Elkin I, Gibbons RD, Shea TM, Sotski SM, et al. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. Journal of Consulting and Clinical Psychology. 1995; 63:841–847. [PubMed: 7593878]

Fava M. Augmentation and combination strategies in treatment-resistant depression. Journal of Clinical Psychiatry. 2001; 62 Suppl 18:4–11. [PubMed: 11575733]

Fava M, Iosifescu DV, Pedrelli P, Baer L. Reliability and validity of the Massachusetts general hospital cognitive and physical functioning questionnaire. Psychotherapy and Psychosomatics. 2009; 78(2):91–97. Epub 2009 Feb 13. [PubMed: 19218827]

Fournier JC, Derubeis RJ, Hollon SD, DImidjian S, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. Journal of the American Medical Association. 2010; 303(1):47–53. [PubMed: 20051569]

Frisch, MB. Manual and treatment guide for the Quality of Life Inventory. Minneapolis, MN: National Computer Systems, Inc; 1994.

Frisch MB, Clark MP, Rouse SV, Rudd MD, et al. Predictive and treatment validity of life satisfaction and the quality of life inventory. Assessment. 2005; 12(1):66–78. [PubMed: 15695744]

Hamilton M. A rating scale for depression. Journal of Neurology, Neurosurgery and Psychiatry. 1960; 23:56–61.

Hollon SD, DeRubeis RJ, Evans M, Weimer MJ, et al. Cognitive therapy and pharmacotherapy for depression: Singly and in combination. Archives of General Psychiatry. 1992; 49:774–781. [PubMed: 1417429]

Joffe RT, Young T, Leavitt AJ, MacQueen G, et al. Number of episodes and antidepressant response in major depression. International Journal of Neuropsychopharmacology. 1999; 2:111–113. [PubMed: 11281978]

Khan A, Leventhal RM, Khan SR, Brown WA. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. Journal of Clinical Psychopharmacology. 2002; 22(1):40–45. [PubMed: 11799341]

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Medicine. 2008; 5(2):e45. [PubMed: 18303940]

Leuchter AF, Lesser IM, Trivedi MH, Rush AJ, et al. An open pilot study of the combination of escitalopram and bupropion-SR for outpatients with major depressive disorder. Journal of Psychiatric Practice. 2008; 14(5):271–280. [PubMed: 18832958]

Levine J, Schooler NR. SAFTEE. a technique for the systematic assessment of side effects in clinical trials. Psychopharmacology Bulletin. 1986; 22(2):343–381. [PubMed: 3774930]

Levine J, Schooler NR. General versus specific inquiry with SAFTEE. Journal of Clinical Psychopharmacology. 1992; 12(6):448. [PubMed: 1474186]

McGrath PJ, Stewart JW, Fava M, Trivedi MH, et al. Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. American Journal of Psychiatry. 2006; 163(9):1531–1541. [PubMed: 16946177]

Mundt JC, Marks IM, Shear MK, Greist JH. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. British Journal of Psychiatry. 2002; 180:461–464. [PubMed: 11983645]

Nemeroff CB, Heim CM, Thase ME, Klein DN, et al. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. Proceedings of the National Acadamy of Sciences USA. 2003; 100(24):14293–14296. Epub 2003 Nov 13.

Papakostas GI, Petersen TJ, Kinrys G, Burns AM, et al. Aripiprazole augmentation of selective serotonin reuptake inhibitors for treatment-resistant major depressive disorder. Journal of Clinical Psychiatry. 2005; 66(10):1326–1330. [PubMed: 16259548]

Reilly MC, Zbrozek As, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics. 1993; 4(5):353–365. [PubMed: 10146874]

Robins, E. Suicide. Roy, A., editor. Baltimore, MD: Williams & Wilkins; 1986. p. 123-133.

Rush AJ, Bernstein IH, Trivedi MH, Carmody TJ, et al. An evaluation of the Quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: a Sequenced Treatment Alternatives to Relieve Depression trial report. Biological Psychiatry. 2006; 59(6):493–501. Epub 2005 Sep 30. [PubMed: 16199008]

Rush AJ, Carmody TJ, Reimitz PE. The Inventory of Depressive Symptomatology (IDS): clinician (IDS-C) and self-report (IDS-SR) ratings of depressive symptoms. International Journal of Methods of Psychiatric Research. 2000; 9:45–59.

Rush AJ, Gullion CM, Basco MR, Jarrett RB, et al. The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychological Medicine. 1996; 26(3):477–486. [PubMed: 8733206]

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biological Psychiatry. 2003; 54(5):573–583. Erratum p. 585. [PubMed: 12946886]

Rush AJ, Trivedi MH, Stewart JW, Nierenberg MD, et al. Combining medications to enhance depression outcomes (CO-MED): Acute and long-term outcomes: A single-blind randomized study. American Journal of Psychiatry. in press.

Rush AJ, Zimmerman M, Wisniewski SR, Fava M, et al. Comorbid psychiatric disorders in depressed outpatients: Demographic and clinical features. Journal of Affective Disorders. 2005; 87(1):43–55. [PubMed: 15894381]

Sangha O, Stucki G, Liang MH, Fossel Ah, et al. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis and Rheumatism. 2003; 49(2):156–163. [PubMed: 12687505]

Tedlow J, Fava M, Uebelacker L, Nierenberg AA, et al. Outcome definitions and predictors in depression. Psychotherapy and Psychosomatics. 1998; 67:226–270. [PubMed: 9693350]

Trivedi, et al. submitted.

Trivedi MH, Daly EJ. Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug and Alcohol Dependence. 2007; 88 Suppl 2:S61–S71. [PubMed: 17320312]

Trivedi MH, Rush AJ, Gaynes BN, Stewart JW, et al. Maximizing the adequacy of medication treatment in controlled trials and clinical practice: STAR*D measurement-based care. Neuropsychopharmacology. 2007; 32(12):2479–2489. [PubMed: 17406651]

Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychological Medicine. 2004; 34(1):73–82. [PubMed: 14971628]

Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. American Journal of Psychiatry. 2006; 163(1):28–40. [PubMed: 16390886]

Vallejo J, Gasto C, Catalan R, Bulbena A, et al. Predictors of antidepressant treatment outcome in melancholia: psychosocial, clinical and biological indicators. Journal of Affective Disorders. 1991; 21:151–162. [PubMed: 1829739]

Van HL, Schoevers RA, Dekker J. Predicting the outcome of antidepressants and psychotherapy for depression: a qualitative, systematic review. Harvard Review of Psychiatry. 2008; 16:225–234. [PubMed: 18661365]

Wisniewski SR, Eng H, Meloro L, Gatt R, et al. Web-based communications and management of a multi-center clinical trial: the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) project. Clinical Trials. 2004; 1(4):387–398. [PubMed: 16279277]

Wisniewski SR, Rush AJ, Balasubramani GK, Trivedi MH, et al. Self-rated global measure of the frequency, intensity, and burden of side effects. Journal of Psychiatric Practice. 2006; 12(2):71–79. [PubMed: 16728903]

Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. Archives of General Psychiatry. 2001a; 58(8):787–794. [PubMed: 11483146]

Zimmerman M, Mattia JI. The Psychiatric Diagnostic Screening Questionnaire: development, reliability and validity. Comprehensive Psychiatry. 2001b; 42(3):175–189. [PubMed: 11349235]

Table 1

Sociodemographic measures by baseline QIDS-SR

Measure	Baseline QIDS-SR$_{16}$				Analyses							
							Bonferroni Correction					
	Mild N=81	Moderate N=238	Severe N=260	Very severe N=67	Test statistic	p-value	Mild vs. Moderate	Mild vs. Severe	Mild vs. Very severe	Moderate vs. Severe	Moderate vs. Very severe	Severe vs. Very severe
	%	%	%	%								
Age					$\chi^2(6)=12.81$	**0.0461**	0.7968	0.0635	0.0497	0.0328	0.0571	0.6818
18–29	14.8	18.1	21.9	25.4								
30–54	56.8	55.0	60.8	61.2								
55–75	28.4	26.9	17.3	13.4								
Sex					$\chi^2(3)=24.73$	<**.0001**	0.1017	**0.0005**	<**.0001**	0.0123	**0.0004**	0.0363
Male	48.1	37.8	27.3	14.9								
Female	51.9	62.2	72.7	85.1								
Race					p<0.01	0.1002						
White	70.1	61.9	73.2	59.1								
Black	27.3	31.2	22.0	34.8								
Other	2.6	6.9	4.8	6.1								
Hispanic	17.3	15.5	15.8	11.9	$\chi^2(3)=0.86$	0.8349						
Employed	59.3	50.0	49.2	32.8	$\chi^2(3)=10.52$	**0.0147**	0.1495	0.1148	**0.0014**	0.8638	0.0128	0.0163
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**								
Age	46.5 (13.4)	44.2 (13.0)	40.8 (12.6)	41.6 (12.4)	F(3, 642)=5.59	**0.0009**	0.1605	**0.0004**	0.0232	**0.0030**	0.1520	0.6179
Education	14.1 (3.0)	13.8 (2.9)	13.6 (3.1)	13.5 (2.8)	F(3, 624)=0.82	0.4804						
Monthly Household income	3598 (6119)	2726 (4096)	2278 (3413)	1781 (2261)	$\chi^2(3)=13.71$	**0.0033**	0.0408	**0.0041**	**0.0008**	0.3208	0.0257	0.0717

Note: Chi-square for continuous measures indicates Kruskal-Wallis test.

Bold indicates significant (p <.05 for uncorrected analyses, p <.0083 for after Bonferroni correction.QIDS-SR$_{16}$ categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

Abbreviations: QIDS-SR$_{16}$, 16 item quick inventory of depressive symptomatology – self-rated.

Table 2

Clinical Features in Relation to Symptom Severity Groups

Measure	Baseline QIDS-SR$_{16}$				Analyses							
							Bonferroni Correction					
	Mild N=81	Moderate N=238	Severe N=260	Very severe N=67	Test statistic	p-value	Mild vs. Moderate	Mild vs. Severe	Mild vs. Very severe	Moderate vs. Severe	Moderate vs. Very severe	Severe vs. Very severe
	%	%	%	%								
Body mass index					$\chi^2(9)=8.02$	0.5317						
Normal/underweight (<25)	27.2	23.9	24.8	33.3								
Overweight (25–29.9)	21.0	30.3	29.8	24.2								
Obese I (30–34.9)	28.4	18.5	19.8	18.2								
Obese II & III (35+)	23.5	27.3	25.6	24.2								
Menopausal*	39.0	31.3	31.7	33.3	$\chi^2(3)=0.97$	0.8079						
Age at first episode <18	32.1	41.4	49.4	52.2	$\chi^2(3)=10.16$	**0.0173**	0.1405	**0.0063**	0.0132	0.0714	0.1127	0.6809
At least 1 prior episode	75.3	75.1	80.3	79.1	$\chi^2(3)=2.28$	0.5171						
Ever attempted suicide		7.7	12.2	15.4	$\chi^2(3)=13.99$	**0.0029**	0.0090	**0.0002**	**0.0003**	0.1007	0.0613	0.4943
Lifetime severity of suicidally					p <0.01	<**.0001**	0.0297	<**.0001**	<**.0001**	<**.0001**	<**.0001**	0.0967
None	52.6	39.5	20.9	13.8								
Thoughts of dying	25.6	28.3	28.7	18.5								
Suicidal thoughts	11.5	9.9	18.9	24.6								
Specific method	9.0	6.4	11.4	12.3								
Plan/gesture		5.6	6.7	9.2								
Preparation	1.3	2.6	1.2	6.2								
Attempt		7.7	12.2	15.4								
Neglected before age 18	21.0	30.7	43.5	49.3	$\chi^2(3)=22.03$	<**.0001**	0.0943	**0.0003**	**0.0003**	**0.0032**	**0.0048**	0.3951
Emotionally abused before age 18	23.5	32.4	47.3	55.2	$\chi^2(3)=27.34$	<**.0001**	0.1316	**0.0001**	<**.0001**	**0.0007**	**0.0006**	0.2477
Physically abused before age 18	9.9	17.2	23.5	26.9	$\chi^2(3)=10.31$	**0.0161**	0.1130	**0.0079**	**0.0069**	0.0851	0.0777	0.5616
Sexually abused before age 18	16.0	16.0	25.4	35.8	$\chi^2(3)=15.85$	**0.0012**	0.9974	0.0821	**0.0057**	0.0105	**0.0004**	0.0881
Abused before age 18	30.9	39.7	54.2	64.2	$\chi^2(3)=26.99$	<**.0001**	0.1578	**0.0002**	<**.0001**	**0.0012**	**0.0004**	0.1433

Measure	Baseline QIDS-SR$_{16}$				Analyses							
							Bonferroni Correction					
	Mild N=81	Moderate N=238	Severe N=260	Very severe N=67	Test statistic	p-value	Mild vs. Moderate	Mild vs. Severe	Mild vs. Very severe	Moderate vs. Severe	Moderate vs. Very severe	Severe vs. Very severe
	%	%	%	%								
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)								
Age at first episode	29.0 (14.1)	26.1 (14.8)	21.0 (12.8)	22.0 (13.2)	$\chi^2(3)=27.86$	**<.0001**	0.0873	**<.0001**	**0.0017**	**0.0002**	0.0323	0.7160
Years since first episode	17.5 (13.9)	18.1 (14.0)	19.8 (13.2)	19.6 (13.9)	$\chi^2(3)=5.39$	0.1455						
Number of prior episodes	4.9 (8.0)	7.5 (17.1)	10.8 (22.5)	10.9 (24.1)	$\chi^2(3)=4.11$	0.2498						
Number of suicide attempts	0.0 (0.0)	0.12 (0.47)	0.41 (1.99)	0.32 (0.95)	$\chi^2(3)=14.36$	**0.0025**	0.0116	**0.0012**	**0.0003**	0.0803	0.0512	0.5152
Age neglected	7.9 (4.2)	7.6 (4.2)	6.9 (4.2)	7.1 (4.8)	F(3, 230)=0.53	0.6617						
Age emotionally abused	9.6 (3.1)	8.1 (4.0)	7.4 (4.1)	7.8 (4.9)	F(3, 248)=1.64	0.1816						
Age physically abused	11.4 (2.8)	7.9 (3.4)	6.8 (3.9)	7.7 (4.3)	F(3, 121)=3.74	**0.0130**	0.0099	**0.0019**	0.0384	0.1425	0.8667	0.3815
Age sexually abused	9.5 (4.0)	9.3 (4.0)	9.2 (4.0)	8.1 (4.4)	F(3, 137)=0.56	0.6422						

Note: Chi-square for continuous measures indicates Kruskal-Wallis test

Bold indicates significant ($p < .05$ for uncorrected analyses, $p < .0083$ for after Bonferroni correction.QIDS-SR$_{16}$ categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

* Denominator is number of women (see Table 1).

Abbreviations: QIDS-SR$_{16}$, 16 item quick inventory of depressive symptomatology – self-rated.

Table 3

Co-morbidity measures by baseline QIDS-SR$_{16}$

Measure	Baseline QIDS-SR$_{16}$				Analyses							
							Bonferroni Correction					
	Mild N=81	Moderate N=238	Severe N=260	Very severe N=67	Test statistic	p-value	Mild vs. Moderate	Mild vs. Severe	Mild vs. Very severe	Moderate vs. Severe	Moderate vs. Very severe	Severe vs. Very severe
	%	%	%	%								
PDSQ												
Agoraphobia	1.2	8.4	12.7	22.4	$\chi^2(3)=19.60$	**0.0002**	0.0246	**0.0027**	**<.0001**	0.1211	**0.0015**	0.0455
Alcohol abuse	9.9	12.6	8.1	10.4	$\chi^2(3)=2.75$	0.4325						
Bulimia	2.5	10.1	16.2	14.9	$\chi^2(3)=12.52$	**0.0058**	0.0305	**0.0013**	**0.0057**	0.0460	0.2660	0.8063
Drug abuse	2.5	3.8	6.2	11.9	p <0.01	0.0503						
Generalized anxiety	3.7	15.5	24.6	37.3	$\chi^2(3)=32.45$	**<.0001**	**0.0054**	**<.0001**	**<.0001**	0.0119	**<.0001**	0.0373
Hypochondriasis	2.5	0.8	6.9	10.4	p <0.01	**<.0001**	0.2677	0.1792	0.0792	**0.0006**	**0.0005**	0.3330
Obsessive-compulsive	8.6	11.3	12.3	17.9	$\chi^2(3)=3.18$	0.3645						
Panic disorder	2.5	6.3	11.2	28.4	$\chi^2(3)=34.00$	**<.0001**	0.2563	0.0176	**<.0001**	0.0567	**<.0001**	**0.0004**
Post-traumatic stress disorder	4.9	8.8	15.4	23.9	$\chi^2(3)=17.04$	**0.0007**	0.2611	0.0143	**0.0008**	0.0257	**0.0009**	0.0998
Social phobia	7.4	22.3	34.6	41.8	$\chi^2(3)=33.21$	**<.0001**	**0.0029**	**<.0001**	**<.0001**	**0.0024**	**0.0014**	0.2755
Somatoform		1.7	5.4	4.5	p <0.01	**0.0244**	0.5756	0.0472	0.0905	0.0270	0.1813	1.0000
Substance abuse	9.9	14.7	12.0	17.9	$\chi^2(3)=2.86$	0.4141						
Number of psychiatric disorders					$\chi^2(12)=63.54$	**<.0001**	**0.0076**	**<.0001**	**<.0001**	**0.0016**	**<.0001**	0.1749
0	70.4	47.9	35.9	23.9								
1	18.5	27.7	23.2	25.4								
2	7.4	10.5	18.5	16.4								
3	2.5	7.6	8.9	10.4								
4+	1.2	6.3	13.5	23.9								

Note: Chi-square for continuous measures indicates Kruskal-Wallis test.

Bold indicates significant (p <.05 for uncorrected analyses, p <.0083 for after Bonferroni correction.

QIDS-SR$_{16}$ categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

Abbreviations: PDSQ, Psychiatric diagnostic screening questionnaire; QIDS-SR16, 16 item quick inventory of depressive symptomatology – self-rated; SCQ, Self-administered comorbidity questionnaire.

Table 4

Presentation measures by baseline QIDS-SR_{16}

Measure	Baseline QIDS-SR_{16}				Analyses							
							Bonferroni Correction					
	Mild N=81	Moderate N=238	Severe N=260	Very severe N=67	Test statistic	p-value	Mild vs. Moderate	Mild vs. Severe	Mild vs. Very severe	Moderate vs. Severe	Moderate vs. Very severe	Severe vs. Very severe
	%	%	%	%								
Clinical setting					$\chi^2(3)=0.53$	0.9124						
Primary	54.3	54.6	51.5	53.7								
Specially	45.7	45.4	48.5	46.3								
Current episode durations 2+ years	46.9	61.2	55.2	55.2	$\chi^2(3)=5.35$	0.1477						
Chronic/recurrent depression					$\chi^2(6)=8.63$	0.1954						
Chronic only	24.7	24.9	19.7	20.9								
Recurrent only	53.1	38.8	44.8	44.8								
Both	22.2	36.3	35.5	34.3								
IDS-C_{30} lethargic depression	29.6	58.8	81.9	91.0	$\chi^2(3)=103.18$	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0708
Anxious features	65.4	67.6	81.5	85.1	$\chi^2(3)=20.23$	**0.0002**	0.7140	**0.0024**	**0.0065**	**0.0004**	**0.0053**	0.4992
Atypical features	3.7	8.0	22.7	25.4	$\chi^2(3)=34.68$	**<.0001**	0.1892	**0.0001**	**0.0001**	**<.0001**	**<.0001**	0.6432
Melancholic features	2.5	9.5	28.2	57.6	$\chi^2(3)=90.45$	**<.0001**	0.0412	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**
IDS-C_{30} sleep disturbance	77.8	88.2	88.8	97.0	$\chi^2(3)=13.43$	**0.0038**	0.0207	0.0116	**0.0007**	0.8306	0.0330	0.0418
CHRT-SR suicidal thoughts/plans	3.7	12.6	16.5	46.3	$\chi^2(3)=55.17$	**<.0001**	0.0231	**0.0032**	**<.0001**	0.2151	**<.0001**	**<.0001**
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)								
Number of prior antidepressants	1.1 (1.5)	1.3 (1.5)	1.8 (1.8)	2.1 (2.2)	$\chi^2(3)=17.67$	**0.0005**	0.0853	**0.0007**	**0.0022**	0.0084	0.0234	0.4906
Number of concomitant medications	3.1 (3.5)	3.3 (3.0)	2.8 (2.5)	2.7 (2.4)	$\chi^2(3)=5.84$	0.1197						
Current episode duration (months)	55.3 (97.1)	68.9 (117)	62.3 (102)	55.6 (86.9)	$\chi^2(3)=2.23$	0.5259						
$HRSD_{17}$	19.9 (3.1)	22.2 (4.1)	25.4 (4.3)	28.4 (4.8)	F(3, 640)=74.04	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**
IDS-C_{30}	27.6 (6.3)	34.7 (6.9)	41.9 (7.3)	48.0 (7.3)	F(3, 642)=148.08	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**

Measure	Baseline QIDS-SR$_{16}$				Analyses							
							Bonferroni Correction					
	Mild N=81	Moderate N=238	Severe N=260	Very severe N=67	Test statistic	p-value	Mild vs. Moderate	Mild vs. Severe	Mild vs. Very severe	Moderate vs. Severe	Moderate vs. Very severe	Severe vs. Very severe
	%	%	%	%								
QIDS-C$_{16}$	11.7 (2.8)	14.7 (2.7)	17.3 (2.6)	19.4 (2.6)	F(3, 642)=147.49	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**
ASRMS	2.0 (2.4)	1.5 (2.3)	1.4 (2.2)	1.6 (2.3)	$\chi^2(3)$=8.10	0.0441	0.0240	**0.0043**	0.0681	0.5087	0.8561	0.8292
CAST-SR irritability	9.5 (3.5)	11.5 (3.5)	13.5 (3.4)	14.8 (3.3)	F(3, 641)=45.68	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**0.0040**
CAST-SR anxiety	4.2 (2.3)	5.8 (2.8)	7.0 (2.9)	7.7 (2.9)	$\chi^2(3)$=74.36	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0896
CAST-SR mania	5.7 (3.1)	3.8 (2.6)	3.0 (2.6)	2.7 (2.9)	$\chi^2(3)$=58.33	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**0.0002**	**0.0008**	0.1730
CAST-SR insomnia	3.7 (2.2)	4.7 (2.2)	5.6 (2.3)	6.1 (2.2)	F(3, 641)=20.35	**<.0001**	**0.0008**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0995
CAST-SR panic	1.5 (1.5)	2.3 (1.9)	3.1 (2.3)	3.8 (2.5)	$\chi^2(3)$=54.86	**<.0001**	**0.0014**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0437
CHRT-SR loneliness	2.5 (1.8)	2.9 (1.8)	3.8 (1.9)	4.6 (2.3)	F(3, 641)=24.47	**.0001**	0.0561	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0150
CHRT-SR despair	2.6 (1.9)	3.7 (2.0)	5.2 (1.9)	5.8 (2.2)	F(3, 642)=58.50	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0186
CHRT-SR ideation	1.0 (1.7)	1.8 (2.5)	2.2 (2.5)	1.4 (3.5)	$\chi^2(3)$=40.32	**<.0001**	0.0299	**0.0002**	**<.0001**	0.0325	**<.0001**	**<.0001**
CHRT-SR total	6.1 (3.6)	8.4 (4.6)	11.2 (4.6)	14.5 (6.6)	F(3, 641)=53.34	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**0.0002**
CPFQ	21.9 (4.4)	25.9 (5.2)	29.9 (5.0)	32.6 (5.2)	F(3, 642)=82.97	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**
QOLI	0.2 (1.6)	−0.8 (1.7)	−1.8 (1.8)	−2.2 (1.9)	F(3, 638)=38.81	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	0.0738
WSAS	17.5 (7.8)	24.8 (8.5)	30.3 (6.6)	33.1 (7.6)	$\chi^2(3)$=178.23	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**

Note: Chi-square for continuous measures indicates Kruskal-Wallis test.

Bold indicates significant (p <.05 for uncorrected analyses, p <.0083 for after Bonferroni correction.

QIDS-SR$_{16}$ categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

Abbreviations: ASRMS, Altman self-rated mania scale; CAST-SR, Concise associated symptoms tracking scale – self-rated; CHRT-SR, Concise health risk tracking scale – self-rated; CPFQ, Cognitive and physical functioning questionnaire; HRSD$_{17}$, 17-item Hamilton rating scale for depression; IDS-C$_{30}$, 30-item inventory of depressive symptomatology – clinician-rated; QIDS-C$_{16}$ -SR$_{16}$, 16-item quick inventory of depressive symptomatology – clinician-rated, -self-rated; QOLI, Quality of life inventory; WSAS, Work and social adjustment scale.

Table 5

Week 12 outcome measures by baseline QIDS-SR_{16}

Measure	Baseline QIDS-SR_{16}				Unadjusted				Adjusted*			
	Mild n=81	Moderate n=238	Severe n=260	Very severe n=67	Moderate vs. Mild	Severe vs. Mild	Very severe vs. mild	p-value	Moderate vs. Mild	Severe vs. Mild	Very severe vs. mild	p-value
	%	%	%	%	OR	OR	OR		OR	OR	OR	
Exited acute phase	23.5	25.6	30.0	29.9	1.156	1.346	1.359	0.7472	1.250	1.283	1.033	0.8344
At least 1 SAE†	3.7	2.9	4.2	6.0								
At least 1 psychiatric SAE†		0.8	1.5	1.5								
Last 2 consecutive QIDS-SR_{16} <6|8	61.7	39.9	32.3	28.4	0.439	0.299	0.287	**0.0001**	0.599	0.617	0.866	0.2653
Percent QIDS-SR_{16} reduction >50%	48.8	49.8	53.3	58.2	1.095	1.181	1.725	0.4308	1.370	1.874	4.236	**0.0285**
31–40	1.3	10.4	19.0	26.6								
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	β	β	β		β	β	β	
Maximum SAFTEE N worsening	9.7 (8.0)	9.3 (5.7)	9.8 (6.2)	9.5 (6.2)	1.039	1.055	1.017	0.9223	1.011	0.994	0.917	0.8430
Last SAFTEE N worsening	5.1 (5.4)	5.0(4.9)	5.5(5.3)	4.7(4.7)	−0.042	0.017	−0.031	0.9377	−0.014	0.012	−0.200	0.6114
Last QIDS-SR_{16}	5.2 (4.1)	7.3 (4.3)	9.5 (5.9)	10.0 (6.5)	1.890	4.181	4.206	**<.0001**	0.735	1.834	0.388	**0.0364**
Percent QIDS-SR_{16} change	−35 (47.2)	−46(31.5)	−47 (32.5)	−55 (28.7)	−11.20	−12.44	−22.48	**0.0019**	−14.79	−20.05	−36.39	**<.0001**[a]
IDS-C_{30} anxiety subscale	1.8 (1.9)	2.4 (2.0)	2.9 (2.3)	3.0 (2.1)	0.273	0.445	0.451	**0.0013**	0.129	0.097	−36.39	0.1922
Last QOLI	1.42 (2.17)	0.58 (2.11)	−0.51 (2.20)	−0.26 (2.80)	−0.918	−1.969	−1.845	**<.0001**	−0.410	−0.906	−0.222	**0.0174**

Note: models assume either OR=1 or β=1 for reference category of the analysis variable.

QIDS-SR_{16} categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

*adjusted for treatment, age, sex, PDSQ, bulimia, OCD, PTSD, atypical and melancholic features, insomnia, CAST irritability and mania, CHRT, CPFQ, and WSAS.

†models are unestimable.

‡an extremely non-normal distribution required binning.

Post-hoc tests indicate the following pairwise differences after correcting for multiple comparison:

[a]Severe vs. None, Very severe vs. Moderate, Very severe vs. severe.

Abbreviations:: QIDS-SR16, 16 item quick inventory of depressive symptomatology – self-rated; QOLI, Quality of life inventory; SAE, Serious adverse event; SAFTEE, Systematic assessment for treatment emergent events; WSAS, Work and social adjustment scale.

Table 6

Week 28 outcome measures by baseline QIDS-SR_{16}

Measure	Baseline QIDS-SR_{16}				Unadjusted				Adjusted*			
	Mild n=81	Moderate n=238	Severe n=260	Very severe n=67	Moderate vs. Mild	Severe vs. Mild	Very severe vs. mild	p-value	Moderate vs. Mild	Severe vs. Mild	Very severe vs. mild	p-value
	%	%	%	%	OR	OR	OR		OR	OR	OR	
Exited acute phase	35.8	36.1	37.7	37.3	1.043	1.092	1.051	0.9899	1.093	0.960	0.654	0.6185
Severe/intolerable	1.3	0.9	4.0	1.6								
At least 1 SAE†	6.2	5.9	6.5	10.4	0.945	0.936	1.358	0.9144	1.053	0.987	0.909	0.9972
At least 1 psychiatric SAE†	1.2	1.7	2.7	4.5								
Last 2 consecutive QIDS-SR_{16} <6|8	63.0	47.1	40.0	34.3	0.544	0.386	0.369	0.0033	0.624	0.591	0.839	0.2882
Percent QIDS-SR_{16} reduction >50%	59.3	58.1	54.9	72.7	1.037	0.828	1.991	0.0670	1.210	1.353	5.548	**0.0020**[a]
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	β	β	β		β	β	β	
Maximum SAFTEE N worsening	10.9 (8.8)	9.9 (6.1)	10.2 (6.4)	10.1 (6.8)	0.978	0.984	0.954	0.9844	0.968	0.955	0.890	0.9040
Last SAFTEE N worsening	5.5 (6.0)	4.6 (5.1)	5.1 (5.3)	4.7 (5.1)	−0.180	−0.154	−0.126	0.6436	−0.132	−0.147	−0.234	0.8006
Last QIDS-SR_{16}	4.6 (4.3)	6.8 (4.5)	9.1 (6.2)	9.0 (6.2)	1.519	2.011	1.892	**<.0001**	1.394	1.605	1.224	**0.0013**[b]
Percent QIDS-SR_{16} change	−43 (46.6)	−49 (33.5)	−50 (33.6)	−60 (27.5)	−6.871	−6.510	−18.28	**0.0282**	−8.412	−12.73	−32.40	**0.0003**[c]
IDS-C_{30} anxiety subscale	1.9 (1.9)	2.3 (2.1)	2.8 (2.3)	2.8 (2.0)	0.177	0.382	0.367	**0.0090**	0.097	0.111	−0.118	0.3213
Last QOLI	1.61 (2.11)	0.79 (2.18)	−0.10 (2.39)	0.10 (2.77)	−0.853	−1.769	−1.604	**<.0001**	−0.471	−0.942	−0.271	0.0234

Note: models assume either OR=1 or β=1 for reference category of the analysis variable.

QIDS-SR_{16} categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

*adjusted for treatment, age, sex, PDSQ, bulimia, OCD, PTSD, atypical and melancholic features, insomnia, CAST irritability and mania, CHRT, CPFQ, and WSAS.

†models are unestimable.

‡an extremely non-normal distribution required binning.

Post-hoc tests indicate the following pairwise differences after correcting for multiple comparison:

[a]Very severe vs. Moderate, Very severe vs. severe.

[b] Moderate vs. Mild.

[c] Very severe vs. Moderate, Very severe vs. Severe.

Abbreviations:; QIDS-SR16, 16 item quick inventory of depressive symptomatology – self-rated; QOLI, Quality of life inventory; SAE, Serious adverse event; SAFTEE, Systematic assessment for treatment emergent events; WSAS, Work and social adjustment scale.

Table 7

Selected outcome measures by baseline QIDS-SR$_{16}$ and treatment

Treatment	Baseline QIDS-SR$_{16}$												
	Mild			Moderate			Severe			Very severe			
	BUP-SR + ESCIT n=32	ESCIT + PBO n=26	VEN-XR + MIRT n=23	BUP-SR + ESCIT n=72	ESCIT + PBO n=94	VEN-XR + MIRT n=72	BUP-SR + ESCIT n=91	ESCIT + PBO n=80	VEN-XR + MIRT n=89	BUP-SR + ESCIT n=22	ESCIT + PBO n=18	VEN-XR + MIRT n=27	p-value*
	%	%	%	%	%	%	%	%	%	%	%	%	
Week 12													
Early termination	25.0	15.4	30.4	31.9	22.3	23.6	33.0	27.5	29.2	36.4	38.9	18.5	0.6327
Last QIDS-SR$_{16}$ <6	71.0	61.5	56.5	36.1	39.4	35.2	30.0	28.8	31.5	18.2	27.8	33.3	0.7915
% QIDS-SR$_{16}$ reduction >50%	48.4	42.3	56.5	51.4	48.9	49.3	50.0	58.8	51.7	63.6	50.0	59.3	0.7724
Week 28													
Early termination	37.5	30.8	39.1	36.1	34.0	38.9	40.7	33.8	38.2	36.4	50.0	29.6	0.7979
Last QIDS-SR$_{16}$ <6	68.8	76.9	60.9	43.7	44.7	45.1	41.6	41.8	31.5	36.4	22.2	42.3	0.5439
% QIDS-SR$_{16}$ reduction >50%	53.1	65.4	60.9	54.9	57.4	62.0	58.4	59.5	47.2	77.3	61.1	76.9	0.3902
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	
Week 12													
Last QIDS-SR$_{16}$	4.6 (3.4)	5.6 (4.7)	5.4 (4.4)	7.2 (4.3)	7.1 (4.1)	7.5 (4.7)	9.8 (5.8)	9.1 (5.7)	9.6 (6.1)	10.0 (5.4)	10.8 (7.2)	9.6 (7.0)	0.9273
% QIDS-SR$_{16}$ reduction	−32 (50.3)	−35 (46.4)	−38 (45.6)	−45 (32.1)	−47 (29.6)	−44 (33.5)	−46 (31.8)	−49 (31.6)	−46 (34.3)	−56 (21.8)	−51 (32.5)	−57 (31.5)	0.9785
Week 28													
Last QIDS-SR$_{16<6}$	4.3 (3.6)	4.6 (4.3)	5.1 (5.2)	7.0 (4.8)	6.6 (4.4)	6.8 (4.5)	8.5 (5.9)	8.7 (6.0)	10.1 (6.5)	8.6 (5.6)	10.3 (6.7)	8.3 (6.4)	0.8903
% QIDS-SR$_{16}$ reduction	−39 (50.0)	−47 (38.3)	−43 (51.5)	−47 (37.0)	−51 (31.8)	−49 (32.2)	−54 (31.8)	−52 (32.9)	−44 (35.5)	−62 (22.6)	−53 (31.2)	−63 (28.6)	0.4937

* p-value associated with the baseline QIDS-SR$_{16}$ by treatment interaction term.

QIDS-SR$_{16}$ categories: mild = 0–10, moderate = 11–15, severe = 16–20, very severe = 21–27.

Abbreviations: BUP-SR, Bupropion-sustained release; ESCIT, escitalopram; FIBSER, Frequency, intensity, and burden of side effects scale; MIRT, Mirtazapine; PBO, placebo; QIDS-SR$_{16}$, 16 item quick inventory of depressive symptomatology – self-rated; VEN-XR, Venlafaxine-extended release.