Differential change on depressive symptom factors with antidepressant medication and cognitive behavior therapy for major depressive disorder

Boadie Dunlop, *Emory University*
Steven Cole, *Emory University*
Charles B. Nemeroff, *University of Miami*
Helen Mayberg, *Emory University*
Wade Craighead, *Emory University*

Journal Title: Journal of Affective Disorders
Volume: Volume 229
Publisher: Elsevier: 12 months | 2018-03-15, Pages 111-119
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1016/j.jad.2017.12.035
Permanent URL: https://pid.emory.edu/ark:/25593/tp1g7

Final published version: http://dx.doi.org/10.1016/j.jad.2017.12.035

HHS Public Access

Author manuscript

J Affect Disord. Author manuscript; available in PMC 2019 March 15.

Published in final edited form as:

J Affect Disord. 2018 March 15; 229: 111–119. doi:10.1016/j.jad.2017.12.035.

Differential Change on Depressive Symptom Factors with Antidepressant Medication and Cognitive Behavior Therapy for Major Depressive Disorder

Boadie W. Dunlop[1,*], **Steven P. Cole**[2], **Charles B. Nemeroff**[3], **Helen S. Mayberg**[1,4], and **W. Edward Craighead**[1,5]

[1]Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA

[2]Research Design Associates, Inc., Yorktown Heights, NY

[3]Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL, USA

[4]Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, GA, USA

[5]Department of Psychology, Emory University, Atlanta, GA, USA

[*]Corresponding Author: Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 12 Executive Park Drive NE, 3rd Floor, Atlanta, GA, 30329. Phone: 404-727-8474; Fax: 404-727-3700. bdunlop@emory.edu.

Contributors

Drs. Dunlop, Nemeroff, Mayberg, and Craighead designed the PReDICT study and wrote the protocol. Dr. Dunlop managed the literature searches and wrote the first draft of the manuscript. Dr. Cole undertook the statistical analyses. All authors contributed to and have approved the final manuscript.

Conflict of Interest Statement

BWD has received research support from Acadia, Assurex, Axsome, Janssen, NIMH, Otsuka, and Takeda.

SPC has received research support from EPA/SARE and has consulted on studies funded by Emory University, Georgia Department of Human Resources, NIH, NIMH, NSF, and the University of Wisconsin.

WEC is a board member of Hugarheill ehf, an Icelandic company dedicated to the prevention of depression, receives book royalties from John Wiley & Sons, is supported by the Mary and John Brock Foundation and the Fuqua Family Foundations, is a consultant to the George West Mental Health Foundation, a member of the Scientific Advisory Board (SAB) of ADAA, and a member of the SAB for AIM for Mental Health Foundation.

CBN received funding from NIH and the Stanley Medical Research Institute. In the past 3 years, he has served as a consultant to Xhale, Takeda, Taisho Pharmaceutical Inc., ITI, Bracket (Clintara), Total Pain Solutions (TPS), Gerson Lehrman Group (GLG) Healthcare & Biomedical Council, Fortress Biotech, Sunovion Pharmaceuticals Inc., Janssen Research & Development LLC, Magstim, Inc., Navitor Pharmaceuticals, Inc, Actify Neurotherapies, and served on the Board of Directors for the American Foundation for Suicide Prevention, Gratitude America, and the Anxiety Disorders Association of America. CBN is a stockholder in Xhale, Celgene, Seattle Genetics, Abbvie, OPKO Health, Inc., Bracket Intermediate Holding Corp., Antares, and serves on the Scientific Advisory Boards of the American Foundation for Suicide Prevention, Brain and Behavior Research Foundation (BBRF), Xhale, Anxiety Disorders Association of America, Skyland Trail, Bracket (Clintara), RiverMend Health LLC, and Laureate Institute for Brain Research, Inc. CBN reports income sources or equity of $10,000 or more from American Psychiatric Publishing, MagStim, Bracket (Clintara), ITI, CME Outfitters, and Takeda, and has patents on the method and devices for transdermal delivery of lithium (*US 6,375,990B1*) and the method of assessing antidepressant drug therapy via transport inhibition of monoamine neurotransmitters by ex vivo assay (*US 7,148,027B2*).

HSM has received consulting fees from St. Jude Medical Neuromodulation and intellectual property licensing fees from St. Jude Medical Neuromodulation.

Abstract

Background—Major depressive disorder (MDD) is a heterogeneous condition and individual patients are likely to be differentially responsive to specific treatments. In an exploratory factor analysis of three rating scales, the Genome-based Therapeutic Drugs for Depression (GENDEP) trial identified three factors that were differentially associated with outcome to nortriptyline and escitalopram. However, this factor analysis has neither been replicated or applied to a psychotherapy treatment.

Methods—We replicated the GENDEP analytic method in the Emory Prediction of Remission to Individual and Combined Treatments (PReDICT) study. The 17-item Hamilton Depression Rating Scale, Montgomery Asberg Depression Rating Scale, and Beck Depression Inventory were administered to 306 MDD patients in the PReDICT study, which randomized previously untreated adults to 12 weeks of treatment with cognitive behavior therapy (CBT), escitalopram, or duloxetine. Utilizing Item Response Theory methodologies, factor scores were derived from the three scales and the efficacy of the three treatments was compared for the identified factor scores.

Results—Four factors were identified: “Despair,” “Mood and Interest,” “Sleep,” and “Appetite.” These factors closely aligned with the factors identified in GENDEP. Compared to CBT, escitalopram and duloxetine produced more rapid but ultimately similar improvement on the Despair and Mood and Interest factors; no significant differences between treatments emerged on the other factors.

Limitations—The scales contained differing numbers of items pertaining to specific depressive symptoms.

Conclusion—The heterogeneity of MDD can be parsed into a consistent factor structure, with the factors showing differential rapidity, but ultimately similar, improvement across treatments.

Keywords

Precision medicine; rating scales; item response theory; biomarker; psychotherapy; escitalopram; duloxetine

INTRODUCTION

Psychiatric treatment guidelines recommend antidepressant medication or an evidence-based psychotherapy as first-line treatment for major depressive disorder (MDD), with both forms of treatment having proven to be equally efficacious across unselected outpatients (APA 2010; Weitz et al., 2015). However, MDD is a highly heterogeneous syndrome, encompassing a variety of different symptom components (Lux and Kendler, 2010). Whether these individual components are differentially responsive to treatments with differing mechanisms of action has recently gained increased interest, but remains relatively unexamined (Stewart and Harkness, 2012; Fournier et al., 2013).

The rating scales used to assess symptom severity in clinical trials of patients with MDD have been subjected to a variety of factor analyses. Studies taking this approach have usually analyzed a single scale, most often the Hamilton Depression Rating Scale (HAMD, Hamilton, 1960) or Beck Depression Inventory (BDI, Beck et al., 1961). Only a few large

studies comparing psychotherapy and antidepressant medication have been subjected to factor analyses to identify symptom profiles associated with differential change between these treatments. The Treatment of Depression Collaborative Research Program (TDCRP) enrolled 250 MDD patients randomly assigned to interpersonal therapy, cognitive behavior therapy (CBT), imipramine, and pill-placebo (Elkin et al., 1989). Separate factor analyses on the 23-item HAMD (which includes items of over-eating and oversleeping) found that imipramine was superior to CBT on the Sleep factor, and superior to IPT and CBT, considered together as "therapy," on the Cognitive-Affective and Somatic factors (Stewart and Harkness, 2012). Similarly, on the BDI factors, imipramine was superior to CBT on the Somatic factor and superior to therapy on the Affective-Performance factor (Stewart and Harkness, 2012). A separate analysis of the TDCRP data examining the endogenous symptoms of depression defined by the Schedule of Affective Disorders and Schizophrenia (Endicott and Spitzer, 1978) found imipramine superior to CBT, IPT, and placebo (Imber et al., 1990). However, the TDCRP conclusions have been considered tenuous due to site-by-treatment interactions not considered in the preceding analyses (Jacobson and Hollon, 1996).

Use of a single scale to conduct factor analyses is constrained by the limited number of items on each questionnaire. The HAMD in particular has been plagued by numerous inconsistent factor analyses, yielding two to eight factors (Bagby et al., 2004). Factor analyses of the BDI have also produced inconsistent factors (Beck et al., 1988). Moreover, self- and clinician-rated scales differ in important ways in measuring aspects of depression (Dunlop et al., 2010, 2014), so incorporating both vantage points provides a more complete view of the illness (Uher et al., 2012). To address these concerns, a multiple-scale factor analysis was performed by the Genome-based Therapeutic Drugs for Depression (GENDEP) investigators (Uher et al., 2008). This integrated factor analysis of the 17-item HAMD, BDI-I, and the Montgomery Asberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979) for 660 patients who participated in the GENDEP trial identified three factors: Observed Mood, Cognitive, and Neurovegetative.

The GENDEP study was a partially-randomized, two-arm trial that compared 12 weeks' treatment with the selective serotonin reuptake inhibitor (SSRI) escitalopram versus the noradrenergic tricyclic antidepressant (TCA) nortriptyline. Using the identified factors, escitalopram proved superior to nortriptyline on the Observed Mood and Cognitive factors, whereas nortriptyline was superior on the Neurovegetative factor (Uher et al., 2009). Despite the great potential clinical relevance of this factor analysis, replication of the three identified factors has not been attempted, nor studied in patients treated with serotonin-norepinephrine reuptake inhibitors (SNRIs) or psychotherapy. Thus, there is need to replicate and extend the GENDEP factor analysis to additional datasets.

The Emory Prediction of Remission to Individual and Combined Treatments (PReDICT) study was a randomized, 3-arm trial comparing the efficacy of escitalopram, duloxetine, and CBT among treatment-naive patients with MDD (Dunlop et al., 2012, 2017). PReDICT employed the same three scales of depression severity as GENDEP: the HAMD-17, BDI, and MADRS. We undertook a factor analysis following the same procedures as the GENDEP study, with the aim of replicating the identified factors and examining whether outcomes across the factors differed by treatment.

METHODS

Study Design

A full description of the PReDICT study has been presented elsewhere (Dunlop et al., 2012). PReDICT was conducted through the Emory University Mood and Anxiety Disorders Program (MAP) and involved two clinics: 1) the primary MAP Clinic at Emory University, including a satellite location in Stockbridge, Georgia, and 2) a purely Spanish-speaking clinic at Grady Hospital in Atlanta (Dunlop et al., 2017). All patients provided written informed consent prior to participating in the study. The study was approved by Emory's Institutional Review Board and the Grady Hospital Research Oversight Committee, and was registered at clinicaltrials.gov as NCT00360399.

Participants

The PReDICT study enrolled adults aged 18–65 years who met DSM-IV-TR criteria for MDD without psychotic features as their primary current psychiatric diagnosis, identified by a Structured Clinical Interview for DSM-IV (First et al., 1995) and an independent psychiatrist's interview. Key eligibility criteria included being treatment naïve, defined as never (lifetime) having received treatment for a mood disorder with either: (i) a marketed antidepressant at a minimum effective dose for 4 or more consecutive weeks; or (ii) 4 or more sessions of an evidence-based and structured psychotherapy (CBT, IPT or behavioral marital therapy). Patients also had to score ≥18 at screening and ≥15 at the baseline visit on the HAMD total score. Key exclusion criteria included: any (lifetime) prior exposure to escitalopram, citalopram, or duloxetine; any medically significant or unstable medication condition that could impact study participation or data interpretation; any current (past 12 months) diagnosis of obsessive compulsive disorder, eating disorder, substance dependence, or dissociative disorder, or for meeting criteria for substance abuse within the 3 months prior to baseline. In addition, any clinically significant suicide or homicide risk, or the presence of psychotic symptoms were exclusionary.

Randomization and Treatments

Once eligibility for randomization was confirmed, patients were assigned in a 1:1:1 ratio to receive 12 weeks of treatment with either escitalopram (10–20 mg/d, end-point mean: 16.2±5.1 mg/d), duloxetine (30–60 mg/d, end-point mean: 48.0±15.0 mg/d), or 16 individual one-hour sessions of manualized CBT (mean sessions attended: 11.4±4.7) (Beck et al., 1979), delivered via twice weekly sessions for the first four weeks, followed by weekly sessions for the subsequent 8 weeks. Concomitant psychoactive medications were not permitted, with the exception of non-benzodiazepine sedative/hypnotics, which could be used up to three times per week.

Measures

The HAMD, BDI-I and MADRS were administered at baseline, then weekly for the first six weeks, and then biweekly for the final six weeks of treatment. A validated Spanish-language version of the BDI-I was used for Spanish-speaking patients (Conde-Lopez et al., 1976). Structured interview guides were used for the HAMD (Williams et al., 1989) and MADRS

(Williams et al., 2008) to standardize prompts and anchors across all raters, with validated Spanish-language versions administered by Spanish-speaking raters. The HAMD-17 consists of 9 items rated 0–4 and 8 items rated 0–2. The MADRS comprises 10 items rated 0–6. The self-report BDI-I contains 21 items scored 0–3. The clinician rated measures, i.e., the HAMD and MADRS, were administered by a rater masked to treatment assignment. Fourteen bachelors-level individuals with at least three months of supervised rater training served as masked raters over the course of the trial. Video-recorded interviews were used to assess interrater reliability. Intraclass correlation coefficients were 0.91 (95% CI: 0.75–0.99) for the HAMD and 0.94 (95% CI:0.81–0.99) for the MADRS.

Statistical Analysis

Approach to analysis—Following the approach of Uher and colleagues (2008) and allowing for assessment of the psychometric properties across the treatment period, we selected one week from each randomized subject to derive a "random week dataset" for use in the factor analyses. When ratings for a selected week were missing for an individual, the closest week's data were used. Using a random week data set avoids problems that arise from using only baseline scores that have a limited distribution due to the minimum severity eligibility criterion, or using only exit scores, which may be biased by having higher scores only in patients proving unresponsive to treatment or by psychological factors related to the end of study participation.

Scale Dimensions and Factors—We first conducted confirmatory factor analysis (CFA) to evaluate the unidimensionality of each scale separately and of all 48 items for the three scales combined. CFA with maximum likelihood estimation (ML) was conducted with IBM SPSS Amos v. 24.0 (Arbuckle, 2014). Goodness of fit of the one-factor CFA models was assessed with indices widely used in the applied literature that have favorable performance in Monte Carlo research (Brown, 2015): standardized root mean square residual (SRMR), an absolute average of the residual correlation matrix; root mean square of approximation (RMSEA); comparative fit index (CFI); and the Tucker-Lewis incremental fit index (TLI). Because fit indices can be affected by various analytic attributes such as sample size and estimation method, absolute fit index cutoffs are not recommended. However, guidelines for reasonably good fit between a CFA theoretical model and observed data have been suggested: for ML estimation, SRMR values .08 or below; RMSEA values .06 or below; and CFI and TLI values .95 or greater (Brown, 2015; Hu and Bentler, 1999), although CFI and TLI values in the range of .90 and .95 may be indicative of an acceptable model fit (Bentler, 1990). Exploratory factor analysis (EFA) studies were conducted on the random week dataset for the combined scales using IBM SPSS v. 24 principal axis factoring with oblique PROMAX rotation. With a factor analysis oblique rotation, the meaning of factors is ascertained from the pattern matrix. For oblique rotations, the factors are allowed to correlate so the loadings and correlations are distinct. The pattern matrix holds the loadings. Each row of the pattern matrix is a regression equation where the standardized observed variable is expressed as a function of the factors. To determine the number of factors to be extracted, we performed a Monte Carlo simulation of normal random samples that parallel the observed data in terms of sample size and number of variables used (O'Connor, 2000). This parallel analysis served as a comparison against the observed

eigenvalues. In EFA, the higher the loading score of an item on a factor, the more the variable is a pure measure of the factor. Factor loading scores can be categorized as follows: >.71 (50% overlapping variance) excellent; .63–.71 very good; .55–.62 good; .45–.54 fair; and .32–44 poor (a loading score of .32 is equivalent to 10% overlapping variance) (Tabachnick and Fidell, 2014). Items that cross-load onto more than one factor are not considered significant if the difference in loading scores is <.2.

Using item response theory (IRT) methods, scale scores were derived for the random week dataset for each factor using a graded response model with IRTPRO v. 4.0 (Cai et al., 2016). The parameters from these analyses were used to generate the IRT scale scores for each factor for baseline, weeks 1 to 6, 8, 10, and 12 from the longitudinal database.

Scale internal consistency was assessed with Cronbach's α and with corrected item-total correlations for HAMD, MADRS, BDI and Factors 1 to 4.

Treatment Effects—Treatment outcomes were analyzed with linear mixed models (LMM), which use all valid data for participants and can withstand some missing data provided the missing data are missing-at-random (MAR). The MAR assumption was evaluated by assessing the relationship between "missing-ness" for a visit and severity of depression on the preceding visit. Missing-ness for a visit was coded as 0 for missing and 1 for valid data. For each of the four factors, Pearson correlation coefficients were computed for missing-ness at weeks 1 to 6, 8, 10, and 12. The LMM analyses used an autoregressive covariance structure with individual random intercepts and slopes, fitted with maximum likelihood. The employed models included baseline factor score as a covariate; three-level treatment condition as a fixed factor; linear function of time; and scores at weeks 1 to 6, 8, 10, and 12 as repeated measures. In these models, with baseline score as a covariate, a significant treatment effect indicates an overall group difference occurring between week 1 and week 12. If a significant group difference was found, pairwise contrasts by week were tested.

RESULTS

Sample Characteristics

The clinical, social, and demographic characteristics of the patients included in the factor analyses are presented in Table 1. There were no significant differences at baseline between any of the listed characteristics.

The mean scores on the three scales for the weeks used in the random week dataset were: HAMD: 12.2; MADRS: 17.0; BDI: 12.1. For purposes of comparison, the 660 patients included in the random week data set from the GENDEP study had means of: HAMD: 13.6; MADRS: 18.6; BDI: 18.7 (Uher et al., 2008). Thus, the clinician-rated scales were very similar (1.4–1.6 point difference) with a more substantial 6.6 point difference in the mean self-report BDI scores.

Summary of Measures

Of the 344 patients randomized in PReDICT, 6 were subsequently found to have met a study exclusion criterion, leaving 338 patients with MDD for this analysis. Three patients lacked baseline BDI data and, due to attrition, an additional 29 patients had no rating scale data after the baseline visit, resulting in 306 patients available for the IRT analyses. Forty-seven (15.4%) of the analyzed cases had post-baseline data missing one or more scales for the randomly selected week, thereby requiring substitution with the nearest week. Week 1 was the substituted week for 17 cases; Week 2 for 10 cases, and all subsequent weeks were substituted for between 1–5 cases.

Using scores from the random-week dataset, internal consistency was calculated using Cronbach's α and was found to be excellent (> .9) for BDI, .91; and good (> .8) for HAMD, .82; and good for MADRS, .86. Recalculating the Cronbach's α after exclusion of the 46 cases with substituted weeks changed the Cronbach's α value by <1% for all scales. Week 12 (completion) scores were available for 248 of the patients. For the HAMD corrected item-total correlations, 9 of the 17 items (52.9%) were low (<0.4) and 4 (23.5%) were adequate (>.5). For MADRS, only 1 of the 10 items was low (10.0%) while 6 (60.0%) were adequate. For BDI, 4 of 21 items (19.1%) were low while 14 (66.7%) items were adequate.

Confirmatory Factor Analyses

Scale unidimensionality was assessed separately with single-factor CFA for HAMD, MADRS, and BDI measures. Goodness-of-fit indices for HAMD were SRMR = .064, RMSEA = .069 (90% CI, .059, .079), CFI = .835, and TLI = .811; only SRMR and marginally RMSEA indicating a good fit; CFI and TLI were less than .90 reflecting a poor fit. For BDI, SRMR = .061, RMSEA = .076 (90% CI, .069, .084), CFI = .857, and TLI = .841; only SRMR indicated a good fit. For MADRS, SRMR = .051, RMSEA = .085 (90% CI, .068, .103), CFI = .930, and TLI = .910; SRMR indicated a good fit, CFI and TLI were in the acceptable range, and only RMSEA indicated a poor fit. Thus, for all three scales the results were mixed, but the evidence was stronger for MADRS unidimensionality than for HAMD or BDI.

To assess the premise that the combined three scales measured a single construct, a single-factor CFA was conducted with all 48 items. Goodness-of-fit indices were SRMR = .088, RMSEA = .111 (90% CI, .108, .114), CFI = .524, and TLI = .503. All indices indicated a poor model fit. With such evidence of more than one underlying dimension, exploratory factor analyses (EFA) were conducted.

Exploratory Factor Analyses

HAMD item 17 (Insight) had a highly limited distribution (99.3% = 0 at baseline) and was very weakly correlated with other items; it was therefore excluded rendering the included items consistent with the GENDEP factor analysis (Uher et al., 2008). The ratio between the first and second eigenvalues was high (4.7) suggesting unidimensionality. To further examine the factorial structure, we performed a Monte-Carlo simulation of normal random samples that parallel the observed data in terms of sample size and number of variables used.

This parallel analysis served as a comparison against the observed eigenvalues. As shown in the scree plot in Supplementary Fig. 1, the first four observed eigenvalues clearly exceed the values generated from the parallel analysis. The differences between the observed and simulated eigenvalues were somewhat smaller for the fifth through eighth factors. Consequently, a four-factor solution was examined.

The items that loaded onto each of the four factors are presented in Table 2, and the factors were labeled Despair, Mood and Interest, Sleep, and Appetite. Values presented in the table are regression coefficients from the pattern matrix. Three items cross-loaded onto two scales. The clinician-rated suicide items (HAMD 3 and MADRS 10) both loaded very heavily on the Despair factor, with the HAMD-3 item negatively cross loading onto the Mood and Interest factor (−.334) and the MADRS 10 item weakly loading onto the Appetite factor (.322). The third cross-loaded item, BDI item 4 (Enjoyment), primarily loaded on the Mood and Interest factor (.455), with weaker loading on Despair (.363).

The factors were weakly to moderately correlated with each other (Supplementary Table 1). The highest correlation was between Despair and the Mood and Interest factors (r=.650) and the lowest was between the Despair and the Appetite factors (r=.267). Table 3 shows the loading of the symptoms on the four PReDICT factors along with the symptom loading on the three factors identified in the GENDEP analysis (Uher et al., 2008).

Cronbach's α was excellent (> .9) for the Mood and Interest, .909, and Despair, .901, factors; and good (> .8) for Sleep, .831. For the Appetite factor, α was acceptable (> .7), .789. For the Despair factor corrected item-total correlations, none of the items were low (<.4) while 13 of the 16 items (87.5%) were acceptable (> .50). For the Mood and Interest factor, only 1 of 16 items (6.3%) was low and 13 (81.3%) were adequate. For the Appetite factor, all 5 items were adequate. For the Sleep factor, 3 of 4 items (75.0%) were adequate.

Impact of Treatment on Factors

The MAR assumption was evaluated by assessing the relationship between missing-ness for a visit and severity of depression on the preceding visit. Missing-ness for a visit was coded as 0 for missing and 1 for valid data. For each of the four factors, Pearson correlation coefficients were computed for missing-ness at weeks 1 to 6, 8, 10, and 12. For example, for the Despair factor, the correlation between missing-ness at week 2 and depression severity at week 1 was $r = .01$, $p = .842$. Nine correlations were computed for each of the four factors. Correlations were small, ranging from −.10 to .17, and of the 36 correlations computed only two were statistically significant, one for the Despair factor and one for Mood and Interest factor. Overall, there was minimal evidence that missing data were associated with severity of depression on a preceding visit.

There were no significant differences between treatment groups on any of the factor scores at baseline. Table 4 and Figure 1 demonstrate the effects of the three treatments on the identified factors. For the **Despair** factor, there were significant effects of time ($p < .001$) and treatment condition ($p < .001$). Univariate tests of weekly scores indicated significant treatment condition effects at weeks 1, 2, 3, and 4 ($p < .001$) and at weeks 6 and 8 ($p < .05$). Pairwise comparisons indicated that the escitalopram and duloxetine groups had

significantly lower scores than the CBT group at weeks 1, 2, 3, 4 ($p < .001$), and at 5 and 6 ($p < .05$). Duloxetine also had a lower mean score than CBT at week 8 ($p < .01$). There were no significant differences in treatment conditions during weeks 10–12.

For the **Mood and Interest** factor there were significant effects of time ($p < .001$) and an overall effect of treatment condition ($p = .008$). Univariate tests of weekly scores indicated significant treatment condition effects at weeks 2 ($p < .01$), 3 ($p < .05$), 4 ($p < .01$), and 8 ($p < .05$). Pairwise comparisons revealed that duloxetine group scores were significantly lower than CBT at weeks 2 ($p <. 001$), 3 ($p <. 05$), 4 ($p <. 001$), 6 ($p < .05$) and 8 ($p < .01$). Escitalopram group scores were significantly lower than CBT at weeks 1 and 2 ($p < .05$) and at week 4 ($p < .01$). At week 8, duloxetine group scores were significantly lower than escitalopram group scores ($p < .05$). There were no significant differences in treatment conditions during weeks 10–12.

For the **Sleep** factor, there was a main effect of time ($p < .001$), but no significant overall treatment differences ($p < .372$) or treatment x time interactions ($p < .600$). Similarly, for the **Appetite** factor, there was a significant main effect of time ($p < .001$), but not of treatment group ($p < .559$) or treatment x time interaction ($p < .303$).

Due to the clinical importance of suicide, we examined an additional dimension which captured each of the three suicide items from each scale (HAMD 3, MADRS 10, and BDI 9). All three items loaded on the Despair factor. As shown in Supplementary Fig. 2, there was a significant main effect of time ($p < .001$), but not of treatment group ($p < .565$) or treatment x time interaction ($p < .315$). Importantly, there was no point within the treatment course where suicidal ideation appeared to increase above baseline among the treated patients, suggesting no significant impact of CBT or either antidepressant on increasing suicidal ideation in this adult sample.

DISCUSSION

This paper examined the factor structure of the combined results of three depression rating scales-- HAMD, MADRS and BDI--among 306 treatment naïve adults randomly assigned to escitalopram, duloxetine, or CBT. Four factors were identified: Despair, Mood and Interest, Sleep, and Appetite, providing evidence that MDD is a heterogeneous disorder with multiple underlying factors. Consistent with the total scores on the scales, which found no significant overall differences between the treatments (Dunlop et al., 2017), there were no significant differences among treatment conditions by the end of the 12 weeks of treatment. There was, however, a faster rate of change for the antidepressants over CBT on the Despair and the Mood and Interest factors, but by weeks 10–12 no significant differences remained among the three treatments.

Three prominent implications result from this factor analysis. First, we identified factors across three scales that closely aligned with those identified in the GENDEP study. This close congruence of factors is remarkable, particularly considering the inconsistencies across numerous factor analyses conducted with the individual scales (Bagby et al., 2004). Both studies identified a primarily despair factor and a primarily mood/interest factor, and one or

two factors consisting primarily of sleep and appetite items. Although Uher and colleagues labelled their first factor as “Cognitive,” the label of “Despair” is preferred because it better reflects the items that loaded significantly on the first factor; the specific items loading on this first factor were largely overlapping in the two studies. The identified factors should be considered particularly robust when considering that they were derived from patient samples from two studies (GENDEP and PReDICT) that differed in several important ways, including: 1) being conducted on different continents in multiple languages; 2) using a racially homogenous sample versus a racially heterogenous sample; 3) using a clinically heterogeneous sample recruited from clinic patients versus a treatment-naïve, primarily advertising-derived research sample; 4) using separate groups of raters without a shared training experience and no inter-rater reliability assessments between the trials; 5) using unblinded versus blinded assessments; 6) permitting use of concomitant benzodiazepines (one-third of GENDEP participants) versus prohibition of benzodiazepines in PReDICT; and 7) substantially differing treatment modalities. Although both studies employed escitalopram, GENDEP’s comparator was a tricyclic antidepressant, nortriptyline, whereas PReDICT’s other arms included duloxetine and a psychotherapy, CBT. These differences in treatments and concomitant medications are particularly relevant, given that random-week data sets across the 12 weeks of treatment were analyzed, which could have allowed for treatment effects to alter the factor loadings. The similarity of the identified factors despite these differences between studies strongly supports their validity.

Despite these similarities, two differences were present between the factors identified in the PReDICT analysis and those of the GENDEP trial. 1) The GENDEP trial identified three factors and PReDICT four. However, the third “Neurovegetative” factor in GENDEP contained the sleep, appetite, and sexual interest items; sleep and appetite separated as two factors in the PReDICT analysis, and sexual interest loaded on the PReDICT Mood and Interest factor. It is possible that these differences emerged from GENDEP’s use of a tricyclic antidepressant, which may have had stronger effects on sleep and appetite, and less impact on sexual function than escitalopram. Indeed, in the GENDEP trial, nortriptyline proved superior to escitalopram on the Neurovegetative factor (Uher et al., 2009). 2) The “Observed Mood” factor in GENDEP differed from the “Mood and Interest” factor in PReDICT primarily in the loading of anxiety items, many of which loaded on the GENDEP factor but not the PReDICT factor. This difference may have stemmed from the differing characteristics of the patients enrolled in the two trials. GENDEP simply required patients to have at least a “moderately severe” major depression. Unlike PReDICT, in GENDEP there was no explicit requirement that MDD be the primary psychiatric diagnosis, and comorbid OCD was not exclusionary. It is therefore possible that differences in anxiety disorder burden between the studies may have contributed to the differential loading of anxiety symptoms.

Second, when compared to CBT the antidepressants produced faster but ultimately the similar magnitude of improvement on the Despair and Mood and Interest factors of depressions. These findings are consistent with previous studies that have evaluated the length of time required for the respective treatments to produce remission evaluated on general depression measures (*cf.* Keller et al., 2000; see Craighead and Dunlop, 2014), indicating that temporal differences in response are likely driven by changes on these two

factors of MDD. That antidepressants effected faster change than did CBT on the Despair and Mood and Interest factors indicates the potential value of identifying specific domains of symptoms across the measures generally used to evaluate outcomes of various treatments for a disorder as heterogeneous as MDD. Because a more rapid response is of paramount importance for patients with higher scores on the Despair factor, one might tentatively conclude the current data support the use of antidepressants or the combination of antidepressants with psychotherapy for this group of patients.

Third, the consistency of the identified factors between the GENDEP and PReDICT analyses suggests that they may be of value for biomarker discovery and mediators of change. The more rapid improvement in the Despair factor with the medications is consistent with the neurocognitive model of antidepressant action (Roiser et al., 2012; Warren et al., 2015). Several studies have demonstrated changes in affective biasing of attention within one week of starting antidepressant medication treatment (Warren et al., 2015), likely mediated in part via the action of medications to reduce reactivity of the amygdala (Sheline et al., 2001; Fu et al., 2004; Anand et al., 2007) and subgenual cingulate cortex (Keedwell et al., 2009) to negative stimuli. CBT has demonstrated similar effects, though these changes have only been observed after longer treatment durations (Fu et al. 2008, Ritchey et al., 2011). Negative cognitions are a prominent target for change in CBT, but several authors have questioned whether improvement in negative schema is a necessary mechanism for recovery from depression with CBT (Ilardi and Craighead, 1999; Whisman, 1999; Kazdin, 2007). Our analysis indicates that both treatments address this group of symptoms, which is consistent with prior studies (Imber et al., 1990; Rector et al., 2000), but that the “bottom-up” engagement of limbic structures by medications acts more rapidly than the “top-down” enhancement of affective control purported to be engaged by CBT (Rosier et al., 2012).

The primary strength of this analysis is that it closely replicated a prior factor analysis using identical methods (Uher et al., 2008). A potential limitation, which applies to both the current work and the GENDEP factor analysis, is that symptoms are covered by differing numbers of items across the scales, which may have impacted the factors identified. For example, there may have been too few items relating to anxiety to identify an anxiety factor. Additionally, although validated Spanish-language versions of the rating scales were used for Spanish-speaking patients, imperfect aspects of translation may have added variability to the factor analysis.

There are some additional points of interest that emerge from this analysis. The contrasting medication treatment outcomes in GENDEP and PReDICT are relevant, in that in PReDICT duloxetine and escitalopram did not separate on any of the factors, whereas in GENDEP escitalopram was superior to nortriptyline on the Observed Mood factor and nortriptyline was superior to escitalopram on neurovegetative factor (Uher et al., 2009). The absence of differences in the speed of improvement on the neurovegetative symptoms of sleep and appetite in our analysis also contrasts with prior findings that TCAs, which have potent anticholinergic and antihistaminergic properties, improve these symptoms faster than psychotherapy (DiMascio et al., 1979; Stewart and Harkness, 2012), thereby demonstating

the importance that specific medications may have in evaluating the effects of antidepressants on depressive factors.

These differences underscore the importance of recognizing that SNRIs and TCAs, despite sharing a mechanism of norepinephrine transporter blockade, should not be considered interchangeable in terms of clinical effects. Also of interest is that neither the GENDEP nor the PReDICT analyses identified an anxiety factor, though this may have stemmed from the limited number of anxiety items in the rating scales. Finally, our evaluation of the suicide dimension found no differences in suicide scores between antidepressant medication and CBT treated patients. Thus, among adults without imminent suicide risk, these treatments neither induced nor aggravated suicidal thoughts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of the Funding Source

This work was supported by the National Institutes of Health (H.S.M., P50 MH077083; W.E.C, RO1MH080880; B.W.D., K23 MH086690; and David S. Stephens, UL1 RR025008, M01 RR0039). Forest Labs and Elli Lilly Inc. donated the study medications, escitalopram and duloxetine, respectively, but were otherwise uninvolved in study design, data collection, data analysis, or interpretation of findings.

References

American Psychiatric Association. Practice Guideline for the Treatment of Patients with Major Depressive Disorder. 3. American Psychiatric Association; Washington D.C: 2010.

Anand A, Li Y, Wang Y, Gardner K, Lowe MJ. Reciprocal effects of antidepressant treatment on activity and connectivity of the mood regulating circuit: an FMRI study. J Neuropsychiatry Clin Neurosci. 2007; 19:274–282. [PubMed: 17827412]

Arbuckle, JL. Amos (Version 24.0) [Computer Program]. Chicago: IBM SPSS; 2014.

Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry. 2004; 161:2163–2177. [PubMed: 15569884]

Beck, AT., Rush, AJ., Shaw, BF., Emery, G. Cognitive Therapy of Depression. Guilford; New York: 1979.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961; 4:561–571. [PubMed: 13688369]

Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev. 1988; 8:77–100.

Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990; 107:238–246. [PubMed: 2320703]

Brown, TA. Confirmatory factor analysis for applied research. Guilford; New York: 2015.

Cai, L., Thissen, D., du Toit, S. IRTPRO (Version 4.) [Computer Program]. Skokie, IL: 2016.

Conde-López V, Chamorro TE, Useros-Serrano E. Critical study of the reliability and validity of Beck's Rating Scale for the measurement of depression. Arch Neurobiol. 1976; 39:313–338.

Craighead WE, Dunlop BW. Combination psychotherapy and antidepressant medication treatment for depression: For whom, when, and how. Ann Rev Psychol. 2014; 65:267–300. [PubMed: 24405361]

DiMascio A, Weissman MM, Prusoff BA, Neu C, Zwilling M, Klerman GL. Differential symptom reduction by drugs and psychotherapy in acute depression. Arch Gen Psychiatry. 1979; 36:1450–1456. [PubMed: 518245]

Dunlop BW, Li T, Kornstein SG, Friedman ES, Rothschild AJ, Ferguson J, Pederson R, Ahmed S, Ninan P, Keller M. Correlation between patient and clinician assessments of depression severity in the PREVENT study. Psychiatry Res. 2010; 177:177–183. [PubMed: 20304503]

Dunlop BW, Binder EB, Cubells JF, Goodman MG, Kelley ME, Kinkead B, Kutner M, Nemeroff CB, Newport DJ, Owens MJ, Pace TWW, Ritchie JC, Aponte-Rivera V, Westen D, Craighead WE, Mayberg HS. Predictors of remission in depression to individual and combined treatments (PReDICT): Study protocol for a randomized controlled trial. Trials. 2012; 13:106. [PubMed: 22776534]

Dunlop BW, Kelley ME, Aponte-Rivera V, Kinkead B, Mletzko-Crowe T, Ritchie JC, Nemeroff CB, Craighead WE, Mayberg HS. Effects of patient preferences on outcomes in the Predictors of Remission in Depression to Individual and Combined Treatments (PReDICT) study, Am. J Psychiatry. 2017; 174:546–556.

Dunlop BW, McCabe B, Eudicone JM, Sheehan JJ, Baker RA. How well do clinicians and patients agree on depression treatment outcomes? Implications for personalized medicine. Hum Psychopharmacol Clin Exp. 2014; 29:528–536.

Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Glass DR, Pilkonis PA, Leber WR, Docherty JP, Fiester SJ, Parloff MB. National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. Arch Gen Psychiatry. 1989; 46:971–982. [PubMed: 2684085]

Endicott J, Spitzer RL. A diagnostic interview. The Schedule for Affective Disorders and Schizophrenia. Arch Gen Psychiatry. 1978; 35:837–844. [PubMed: 678037]

First, MB., Spitzer, RL., Gibbon, M., Williams, JBW. Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P, Version 2.0). Biometrics Research Department, New York State Psychiatric Institute; New York: 1995.

Fournier JC, DeRubeis RJ, Hollon SD, Gallop R, Shelton RC, Amsterdam JD. Differential change in specific depressive symptoms during antidepressant medication or cognitive therapy. Behav Res Ther. 2013; 51:392–398. [PubMed: 23644038]

Fu CH, Williams SC, Cleare AJ, Brammer MJ, Walsh ND, Kim J, et al. Attenuation of the neural response to sad faces in major depression by antidepressant treatment: a prospective, event-related functional magnetic resonance imaging study. Arch Gen Psychiatry. 2004; 61:877–889. [PubMed: 15351766]

Fu CH, Williams SC, Cleare AJ, Scott J, Mitterschiffthaler MT, Walsh ND, et al. Neural responses to sad facial expressions in major depression following cognitive behavioral therapy. Biol Psychiatry. 2008; 64:505–512. [PubMed: 18550030]

Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960; 23:56–62. [PubMed: 14399272]

Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model. 1999; 6:1–55.

IBM SPSS Statistics for Windows (Version 24.0) [Computer Program]. Armonk, NY: IBM Corp;

Ilardi SS, Craighead WE. The relationship between personality pathology and dysfunctional cognitions In previously depressed adults. J Abn Psychol. 1999; 108:51–57.

Jacobson NS, Hollon SD. Cognitive-behavior therapy versus pharmacotherapy: Now that the jury's returned its verdict, it's time to present the rest of the evidence. J Consult Clin Psychol. 1996; 64:74–80. [PubMed: 8907086]

Kazdin AE. Mediators and mechanisms of change in psychotherapy research. Annu Rev Clin Psychol. 2007; 3:1–27. [PubMed: 17716046]

Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ, Markowitz JC, Nemeroff CB, Russell JM, Thase ME, Trivedi MH, Zajecka J. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. N Engl J Med. 2000; 342:1462–1470. [PubMed: 10816183]

Imber SD, Pilkonis PA, Sotsky SM, Elkin I, Watkins JT, Collins JF, Shea MT, Leber WR, Glass DR. Mode-specific effects among three treatments for depression. J Consult Clin Psychol. 1990; 58:352–359. [PubMed: 2195085]

Lux V, Kendler KS. Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. Psychol Med. 2010; 40:1679–1690. [PubMed: 20059797]

Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979; 134:382–389. [PubMed: 444788]

O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis of Velicer's MAP test. Behav Res Methods Instrum Comput. 2000; 32:396–402. [PubMed: 11029811]

Rector NA, Bagby RM, Segal ZV, Joffe RT, Levitt A. Self-criticism and dependency in depressed patients treated with cognitive therapy or pharmacotherapy. Cog Ther Res. 2000; 24:571–584.

Ritchey M, Dolcos F, Eddington KM, Strauman TJ, Cabeza R. Neural correlates of emotional processing in depression: changes with cognitive behavioral therapy and predictors of treatment response. J Psychiatr Res. 2011; 45:577–587. [PubMed: 20934190]

Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. Neuropsychopharmacol Rev. 2012; 37:117–136.

Sheline YI, Barch DM, Donnelly JM, Ollinger JM, Snyder AZ, Mintun MA. Increased amygdala response to masked emotional faces in depressed subjects resolves with antidepressant treatment: an fMRI study. Biol Psychiatry. 2001; 50:651–658. [PubMed: 11704071]

Stewart JG, Harkness KL. Symptom specificity in the acute treatment of major depressive disorder: A re-analysis of the treatment of depression collaborative research program. J Affect Disord. 2012; 137:87–97. [PubMed: 22252094]

Tabachnick, BG., Fidell, LS. Using multivariate statistics. Pearson; London: 2014.

Uher R, Farmer A, Maier W, Rietschel M, Hauser J, Marusic A, Mors O, Elkin A, Williamson RJ, Schmael C, Henigsberg N, Perez J, Mendlewicz J, Janzing JG, Zobel A, Skibinska M, Kozel D, Stamp AS, Bajs M, Placentino A, Barreto M, McGuffin P, Aitchison KJ. Measuring depression: comparison and integration of three scales in the GENDEP study. Psychol Med. 2008; 38:289–300. [PubMed: 17922940]

Uher R, Maier W, Hauser J, Marusic A, Schmael C, Mors O, Henigsberg N, Souery D, Placentino A, Rietschel M, Zobel A, Dmitrzak-Weglarz M, Petrovic A, Jorgensen L, Kalember P, Giovannini C, Barreto M, Elkin A, Landau S, Farmer A, Aitchison KJ, McGuffin P. Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. Br J Psychiatry. 2009; 194:252–259. [PubMed: 19252156]

Uher R, Perlis RH, Placentino A, Dernovšek MZ, Henigsberg N, Mors O, Maier W, McGuffin P, Farmer A. Self-report and clinician-rated measures of depression severity: can one replace the other? Depress Anxiety. 2012; 29:1043–1049. [PubMed: 22933451]

Weitz E, Hollon SD, Twisk J, van Straten A, David D, DeRubeis RJ, Dimidjian S, Dunlop BW, Faramarzi M, Hegerl U, Jarrett RB, Kheirkhah F, Kennedy SJ, Mergl R, Miranda J, Mohr DC, Rush AJ, Segal ZV, Siddique J, Simmons AD, Vittengl JR, Cuijpers P. Does baseline depression severity moderate outcomes between CBT and pharmacotherapy? An individual participant data meta-analysis. JAMA Psychiatry. 2015; 72:1102–1109. [PubMed: 26397232]

Warren MB, Pringle A, Harmer CJ. A neurocognitive model for understanding treatment action in depression. Phil Trans R Soc B. 2015; 370(1677):20140213. [PubMed: 26240428]

Whisman MA. The importance of the cognitive theory of change in cognitive therapy of depression. Clin Psychol Sci Pract. 1999; 6:300–304.

Williams JB. A structured interview guide for the Hamilton Depression Rating. Arch Gen Psychiatry. 1989; 45:742–747. Spanish translation, Jan 2009, obtained from author.

Williams JBW, Kobak KA. Development and Reliability of the SIGMA: A structured interview guide for the Montgomery-Asberg Depression Rating Scale (MADRS). Br J Psychiatry. 2008; 192:52–58. Spanish translation, Jan 2009, obtained from author. [PubMed: 18174510]

HIGHLIGHTS

- Factor analysis of 3 scales found 4 factors: “Despair”, “Mood and Interest”, “Sleep” & “Appetite.”
- Medication led to faster improvement than CBT only on the Despair and Mood and Interest factors.
- The replication of a prior factor analysis validates this structure of depression heterogeneity.

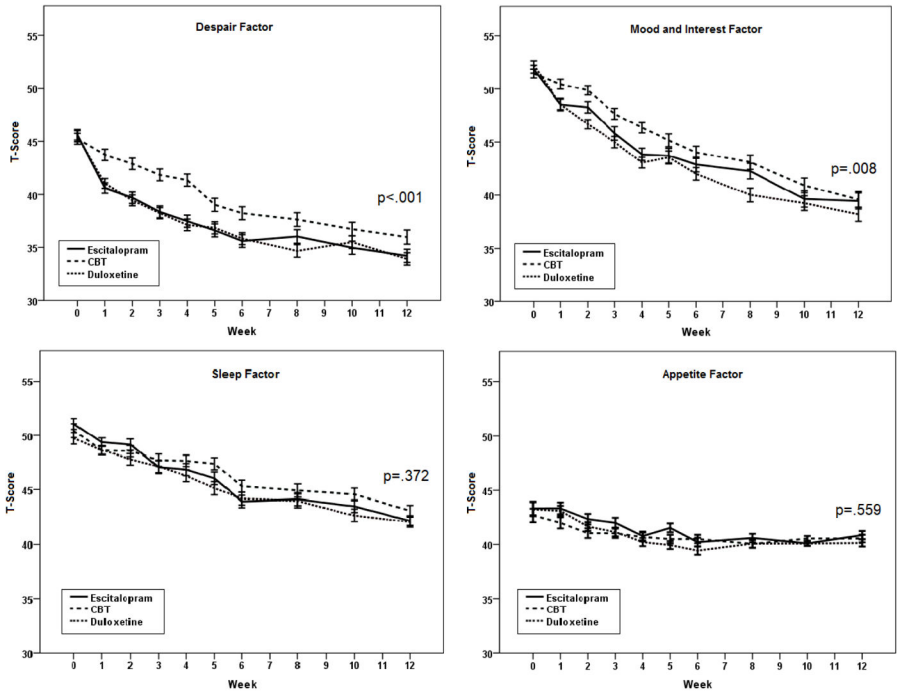

Figure 1.

Differential treatment effects on the four depression factors. Patterns of change in T-scores over time across the three PReDICT treatments for each of the four depression factors. Error bars reflect + 1 standard error.p-values reflect significance of the *F*-statistic from the linear mixed models testing overall treatment condition effects for each factor.

J Affect Disord. Author manuscript; available in PMC 2019 March 15.

Table 1

Sociodemographic and clinical characteristics of sample at baseline

Characteristic	All Patients (n=306)		CBT (n=100)		Escitalopram (n=103)		Duloxetine (n=103)			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	F	p
Age (yrs)	36.67	11.48	39.10	10.71	41.55	12.06	38.33	11.45	2.23	.11
HDRS	19.56	3.61	19.37	3.51	20.00	3.59	19.32	3.69	1.09	.34
MADRS	26.73	5.01	26.68	5.23	26.88	4.78	26.61	5.01	0.08	.92
BDI	22.67	7.17	21.71	7.12	23.04	7.06	23.24	7.39	1.37	.26
	N	%	N	%	N	%	N	%	X^2	P
Sex									0.31	.86
Male	130	42.5	41	41.0	46	44.7	43	41.7		
Female	176	57.5	59	59.0	57	55.3	60	58.3		
Race									6.18	.19
White	150	49.0	52	52.0	45	43.7	53	51.5		
Black	51	16.7	10	10.0	21	20.4	20	19.4		
Other	105	34.3	38	38.0	37	35.9	30	29.1		
Ethnicity									0.32	.85
Hispanic	92	30.1	30	30.0	33	32.0	29	28.2		
Non-Hispanic	214	69.9	70	70.0	70	68.0	74	71.8		
Highest Education Level									5.82	.67
Less than High School	37	12.1	17	17.0	9	8.7	11	10.7		
High School Graduate	51	16.7	13	13.0	18	17.5	20	19.4		
Some College	89	29.1	25	25.0	33	32.0	31	29.1		
Bachelor’s Degree	82	26.8	28	28.0	27	26.2	27	26.8		
Graduate Degree	47	15.4	17	17.0	17	15.5	14	13.6		
Current Anxiety Disorder									0.06	.97
Yes	125	40.8	40	40.0	43	41.7	43	41.7		
No	181	59.2	60	60.0	60	58.3	60	58.3		
Previous Episodes									6.33	.18
1	158	55.1	56	60.2	57	60.6	45	45.0		
2	57	19.9	16	17.2	17	18.1	24	24.0		

Characteristic	All Patients (n=306)		CBT (n=100)		Escitalopram (n=103)		Duloxetine (n=103)			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	F	p
≥3	72	25.1	21	22.6	20	21.3	31	31.0		
Chronic Episode (≥ 2 yrs)									0.62	.73
Yes	92	30.7	31	32.0	28	27.7	33	32.4		
No	208	69.3	66	68.0	73	72.3	69	67.3		

BDI: Beck Depression Inventory; CBT: Cognitive behavior therapy; HAMD: Hamilton Depression Rating Scale, 17-item version; MADRS: Montgomery Asberg Depression Rating Scale

Table 2

Factor loadings across the scales in the PReDICT study

Item	Scale Item No.	Factor 1 Despair	Factor 2 Mood and Interest	Factor 3 Sleep	Factor 4 Appetite
Mood observed	MADRS 1	.226	**.344**	.213	.068
Mood reported	MADRS 2	.206	**.562**	.120	.013
Tension	MADRS 3	**.332**	.138	.301	.040
Sleep	MADRS 4	−.071	−.033	**1.011**	−.044
Appetite	MADRS 5	−.097	.198	−.004	**.857**
Concentration	MADRS 6	−.125	**.555**	.093	.064
Lassitude	MADRS 7	−.031	**.788**	−.089	.043
Inability to feel	MADRS 8	−.029	**.763**	.050	.105
Pessimism	MADRS 9	**.542**	.130	.118	.014
Suicide	MADRS 10	**.745**	−.303	.001	.322
Mood	HAMD 1	.196	**.577**	.087	.007
Guilt	HAMD 2	**.623**	.064	.073	−.005
Suicide	HAMD 3	**.800**	−.334	−.043	.310
Sleep, early	HAMD 4	−.009	.059	**.591**	−.066
Sleep, middle	HAMD 5	.019	−.030	**.697**	.033
Sleep, late	HAMD 6	.019	−.182	**.670**	.079
Activity	HAMD 7	−.013	**.778**	.071	.114
Retardation	HAMD 8	.121	.115	.052	.155
Agitation	HAMD 9	.028	−.041	.128	.100
Anxiety, psychic	HAMD 10	.287	.202	.287	.011
Anxiety, somatic	HAMD 11	−.117	**.396**	.017	.156
Appetite	HAMD 12	−.107	.198	−.010	**.858**
Somatic symptoms	HAMD 13	−.075	**.660**	.106	−.034
Sexual	HAMD 14	−.198	**.526**	−.019	.142
Hypochondriasis	HAMD 15	.047	.222	−.013	−.019
Weight loss	HAMD 16	−.030	−.023	.048	**.421**
Sadness	BDI 1	**.513**	.247	−.001	.006

Item	Scale Item No.	Factor 1 Despair	Factor 2 Mood and Interest	Factor 3 Sleep	Factor 4 Appetite
Future	BDI 2	**.683**	.052	−.017	−.017
Failure	BDI 3	**.747**	.026	−.043	−.078
Enjoyment	BDI 4	.363	**.455**	−.101	.079
Guilt	BDI 5	**.691**	.059	−.005	−.032
Punished	BDI 6	**.696**	−.113	.009	−.067
Disappointed	BDI 7	**.722**	.091	−.020	−.105
Blame self	BDI 8	**.678**	.074	.025	−.154
Suicide	BDI 9	**.646**	−.116	−.050	.209
Crying	BDI 10	**.350**	.153	.086	−.085
Irritable	BDI 11	.324	.318	−.022	−.132
Interest in people	BDI 12	.202	.677	−.164	−.127
Decisions	BDI 13	.187	.569	−.241	.110
Ugly	BDI 14	**.436**	.254	−.114	−.228
Work	BDI 15	.088	**.661**	−.051	−.070
Sleep	BDI 16	−.053	.165	**.600**	.046
Tired	BDI 17	.200	**.552**	.001	−.063
Appetite	BDI 18	.084	.129	.016	**.560**
Weight loss	BDI 19	.100	−.075	.178	.144
Health worry	BDI 20	.274	.211	.010	.005
Sexual interest	BDI 21	−.035	**.524**	−.083	.096

Table 3

Alignment of the PReDICT and GENDEP factors

Despair Factor				Mood and Interest Factor				Sleep Factor				Appetite Factor			
Item name	**Scale item No.**	**PReDICT Loading**	**GENDEP Cognitive Loading**	**Item name**	**Scale item No.**	**PReDICT Loading**	**GENDEP Observed Mood Loading**	**Item name**	**Scale item No.**	**PReDICT Loading**	**GENDEP Neuro-vegetative Loading**	**Item name**	**Scale item No.**	**PReDICT Loading**	**GENDEP Neuro-vegetative Loading**
Suicide	HAM 3	.800	.90	Lassitude	MAD 7	.788	.62	Sleep	MAD 4	1.011	.77	Appetite	HAM 12	.858	.93
Failure	BDI 3	.747	.84	Activity	HAM 7	.778	.56	Sleep, middle	HAM 5	.697	.61	Appetite	MAD 5	.857	.97
Suicide	MAD 10	.745	.64	Inability to Feel	MAD 8	.763	.52	Sleep, late	HAM 6	.670	.71	Appetite	BDI 18	.560	.80
Disappointed	BDI 7	.722	.89	Interest in People	BDI 12	.677	-- (C)	Sleep	BDI 16	.600	.64	Weight loss	HAM 16	.421	.62
Punished	BDI 6	.696	.67	Work	BDI 15	.661	.38	Sleep, early	HAM 4	.591	.52				
Guilt	BDI 5	.691	.81	Somatic symptoms	HAM 13	.660	.50								
Future	BDI 2	.683	.62	Mood	HAM 1	.577	.60								
Blame Self	BDI 8	.678	.85	Decisions	BDI 13	.569	-- (C)								
Suicide	BDI 9	.646	.81	Mood Reported	MAD 2	.562	.68								
Guilt	HAM 2	.623	.72	Concentration	MAD 6	.555	.47								
Pessimism	MAD 8	.542	.86	Tired	BDI 17	.552	-- (C)								
Sadness	BDI 1	.513	.49	Sexual	HAM 14	.526	-- (N)								
Ugly	BDI 14	.436	.60	Sexual Interest	BDI 21	.524	-- (N)								
Crying	BDI 10	.350	.46	Enjoyment	BDI 4	.455	--(C)								
Tension	MAD 3	.332	--	Anxiety somatic	HAM 11	.396	.77								
Irritable	BDI 11	.324	.41	Mood observed	MAD 1	.344	.72								

(C): Item loads on GENDEP Cognitive Factor; (N): Item loads on GENDEP Neurovegetative Factor BDI: Beck Depression Inventory; HAM: Hamilton Depression Rating Scale; MAD: Montgomery Asberg Depression Rating Scale

Table 4

Effects of treatments on depression scale total scores and depression factors

	CBT		Escitalopram		Duloxetine		
	Mean	**95% CI**	**Mean**	**95% CI**	**Mean**	**95% CI**	***p***
Full Scales							
HAMD	12.07	11.19, 12.94	11.07	10.20, 11.94	10.72	9.87, 11.57	.082
MADRS	16.82	15.50, 18.14	15.13	13.83, 16.43	14.36	13.08, 15.64	.029
BDI	10.91	9.79, 12.03	8.84	7.69, 9.99	8.67	7.50, 9.84	.010
Factors							
Despair	−1.15	−1.24, −1.05	−1.39	−1.48, −1.29	−1.39	−1.49, −1.30	<.001
Mood and Interest	−0.56	−0.67, −0.46	−0.70	−0.80, −0.61	−0.78	−0.87, −0.68	.008
Sleep	−0.49	−0.59, −0.39	−0.55	−0.64, −0.46	−0.59	−0.68, −0.50	.371
Appetite	−1.02	−1.09, −0.95	−0.97	−1.04, −0.90	−1.00	−1.07, −0.93	.558

Factor values are IRT-based scores ($M = 0$; $SD = 1$). More negative values indicate less severe symptoms. Means are estimated marginal means across the 9 time points, controlling for baseline. CBT: Cognitive behavior therapy