Journal of Consulting and Clinical Psychology
2012, Vol. 80, No. 6, 995–1006

0022-006X/12/$12.00 DOI: 10.1037/a0030452

Comparative Effectiveness of Medication Versus Cognitive-Behavioral Therapy in a Randomized Controlled Trial of Low-Income Young Minority Women With Depression

Juned Siddique
Northwestern University

Joyce Y. Chung
National Institute of Mental Health, Bethesda, Maryland

C. Hendricks Brown
University of Miami

Jeanne Miranda
University of California, Los Angeles

Objective: To examine whether there are latent trajectory classes in response to treatment and whether they moderate the effects of medication versus psychotherapy. ***Method:*** Data come from a 1-year randomized controlled trial of 267 low-income, young (M = 29 years), minority (44% Black, 50% Latina, 6% White) women with current major depression randomized to antidepressants, cognitive-behavioral therapy (CBT), or referral to community mental health services. Growth mixture modeling was used to determine whether there were differential effects of medication versus CBT. Depression was measured via the Hamilton Depression Rating Scale (Hamilton, 1960). ***Results:*** We identified 2 latent trajectory classes. The first was characterized by severe depression at baseline. At 6 months, mean depression scores for the medication and CBT groups in this class were 13.9 and 14.9, respectively (difference not significant). At 12 months, mean depression scores were 16.4 and 11.0, respectively (p for difference = .04). The second class was characterized by moderate depression and anxiety at baseline. At 6 months, mean depression scores for the medication and CBT groups were 4.4 and 6.8, respectively (p for difference = .03). At 12 months, the mean depression scores were 7.1 and 7.8, respectively, and the difference was no longer significant. ***Conclusions:*** Among depressed women with moderate baseline depression and anxiety, medication was superior to CBT at 6 months, but the difference was not sustained at 1 year. Among women with severe depression, there was no significant treatment group difference at 6 months, but CBT was superior to medication at 1 year.

Keywords: personalized medicine, paroxetine, buproprion, CBT, growth mixture model

Major depression, a disorder with early onset and an often chronic course, imposes a high individual burden of pain, suffering, and disability. Ethnic minority and poor individuals are less likely to receive treatment, particularly guideline-informed care, for major depressive disorder than are White and middle-class individuals (U.S. Department of Health and Human Services [DHHS], 2001). This may be related to the fact that most depression treatment studies include primarily White and middle-class populations (DHHS, 2001), so that little is known about the usefulness of established treatments for more disadvantaged populations. Establishing the effectiveness of depression care in this population is particularly important because rates of depression are elevated in women, younger age cohorts, and those living in or near poverty (Andrade et al., 2003). Because low-income women with depression have few resources and many challenges to overcome to begin and continue with treatment, it is important to make thoughtful, personalized decisions regarding the most effective intervention for a given patient. If an initial treatment strategy is

This article was published Online First October 22, 2012.

Juned Siddique, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine; Joyce Y. Chung, National Institute of Mental Health, Bethesda, Maryland; C. Hendricks Brown, Department of Epidemiology and Public Health, University of Miami Miller School of Medicine; Jeanne Miranda, Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles.

This work was supported by Agency for Healthcare Research and Quality Grant R03-HS018815, National Institute of Mental Health (NIMH) Grant R01-MH040859, and National Cancer Institute Grant K07-CA154862. None of the funding agencies played a role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the article. Juned Siddique, Joyce Y. Chung, and Jeanne Miranda do not have any conflicts of interest to disclose. C. Hendricks Brown is principal investigator on two NIMH grants that evaluate the impact of antidepressants (R01-MH040859 and R01-MH080122). He has also served as a principal investigator on a research grant funded by JDS Pharmaceuticals that supported the evaluation of a behavioral prevention program for suicide.

Correspondence concerning this article should be addressed to Juned Siddique, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 North Lake Shore Drive, Suite 1400, Chicago, IL 60611. E-mail: siddique@northwestern.edu

not effective, patients may not have the additional resources or the desire to pursue another course of treatment. In this article we describe an exploratory analysis to investigate whether there are latent trajectory classes in response to treatment and whether these latent classes moderate the effects of antidepressants versus cognitive behavioral therapy (CBT) in a sample of low-income young minority women.

Comparative effectiveness research has recently received a considerable amount of attention due to the desire by many stakeholders to have more evidence about the relative merits and costs of medical interventions. The U.S. Congress asked the Institute of Medicine as part of the American Recovery and Reinvestment Act of 2009 to determine national priorities for comparative effectiveness research. Among the 100 highest priority research topics identified by the Institute of Medicine, Committee on Comparative Effectiveness Research Prioritization (2009), was "Compare the effectiveness of pharmacologic treatment and behavioral interventions in managing major depressive disorders in adolescents and adults in diverse treatment settings" (p. 111).

There are three key features of comparative effectiveness research: (1) direct comparison of effective interventions, (2) their study under real-world conditions, and (3) research on what patients benefit the most from a given intervention (Sox & Greenfield, 2009; Wang, Ulbricht, & Schoenbaum, 2009). The need for comparative effectiveness research is particularly pressing in the area of mental disorders because only about one fourth of individuals with a mental disorder receive minimally adequate treatment (Wang, Demler, & Kessler, 2002).

Effective treatments for major depression include antidepressant medications and psychotherapies (American Psychiatric Association [APA], 2000; Thase & Kupfer, 1996). Most U.S. psychiatrists favor selective serotonin reuptake inhibitors for first-line medication treatment (Olfson & Klerman, 1993), with treatment extended to at least 6 months to maintain clinical effectiveness (Agency for Healthcare Research and Quality [AHRQ], 1993). CBT is also an effective treatment for major depression. Several studies have found the effectiveness of psychological and medical interventions for depression to be similar (Bortolotti, Menchetti, Bellini, Montaguti, & Berardi, 2008; Casacalenda, Perry, & Looper, 2002; DeRubeis et al., 2005).

Other work has shown that CBT produces sustained clinical gains compared with antidepressant medications that are withdrawn after clinical response (Blackburn, Eunson, & Bishop, 1986; Evans et al., 1992; Kovacs, Rush, Beck, & Hollon, 1981; Miller, Norman, & Keitner, 1989; Shea et al., 1992; Simons, Murphy, Levine, & Wetzel, 1986). In a study of responders to 16 weeks of treatment, patients treated with cognitive therapy were more likely to have a sustained response during 12-month follow-up than were those withdrawn from medications; and they were just as likely to have sustained response as patients who kept taking medications through the follow-up (Hollon et al., 2005). These results suggest that CBT may have important advantages over the long term by preventing relapse after treatment has ended.

Here we performed a comparison of antidepressants versus CBT over the course of a year using data from the Women Entering Care (WECare) study—a clinical trial of predominantly poor young minority women with depression. Initial analyses of the WECare data examined the effectiveness of medication or CBT interventions versus community referral (Miranda et al., 2003, 2006). The WECare investigators found that both guideline-concordant antidepressant medication and a cognitive-behavioral psychotherapy were significantly more effective than referral to mental health care in the community for lowering depressive symptoms and improving functioning at 6 and 12 months after depression was identified. At 6 months, depression treatment outcomes showed that 44.4% of medication, 32.2% of psychotherapy, and 28.1% of community referral patients had remitted (Miranda et al., 2003). At 12-months, remission rates were 51% for medication, 57% for CBT, and 37% for treatment as usual (TAU; Miranda et al., 2006). Unlike the present analysis, these earlier analyses assumed that all participants' trajectories centered around a single average trajectory over time, an assumption that may not be reasonable in the presence of large amounts of between-subjects heterogeneity. In our study, we investigated whether a single underlying trajectory pattern is a valid assumption or whether a more complex model with multiple trajectories fits the WECare data better.

Patients, practitioners, and third-party payers seek guidance as to the type, amount, and cost of treatments that are effective for depression. The current state of the field is that there is no good method to predict which patients with depression will do better on medications versus psychotherapy and, within each treatment modality, which agent or approach is more effective. For most people with depression, the current evidence base does not point to either medication or psychotherapy as working better than the other.

In the present study, we compared the two active WECare interventions (medication and CBT) using a novel statistical method, growth mixture modeling (B. Muthén et al., 2002; B. Muthén & Shedden, 1999), which allowed us to identify and predict multiple response trajectories. We began by identifying several subtypes of clinical response trajectories among the WECare subjects and then compared the effectiveness of antidepressant medication versus cognitive-behavioral therapy within these trajectories.

After modeling the various response trajectories in the WECare data, we classified participants into the response trajectories in terms of their baseline characteristics to identify which patients were more likely to benefit from a given intervention. Our overall goal was to contribute to the development of personalized interventions for individuals with depression.

Method

Study Design

The data used in this analysis come from the WECare clinical trial conducted by Miranda et al. (see Miranda et al., 2003, 2006, for details on their design and methods). Details about participant selection, exclusion, and randomization are summarized the Appendix. Briefly, the study used the Primary Care Evaluation of Mental Disorders (Spitzer et al., 1994) as a depression screen in women attending social service agencies and safety net health clinics (e.g., Title X family planning clinics) in Prince George's and Montgomery Counties, Maryland, and in Arlington and Alexandria, Virginia. Women who screened positive for major depression (11% of those assessed) were invited to participate in confirmatory psychiatric diagnostic telephone interviews. Subjects were excluded if they failed to meet a Composite International Diag-

nostic Interview (CIDI; World Health Organization [WHO], 1997) diagnosis of major depression; were bereaved; were suicidal; had symptoms of mania, psychosis, current alcohol, or other substance abuse; were pregnant or planned to become pregnant; were currently breastfeeding; or were currently receiving mental health care. Those women with confirmed major depressive disorder diagnoses who were willing to participate in the study were randomized to receive pharmacotherapy, CBT, or community referral. Raters were blinded to treatment assignment. The study recruited a diverse ethnic sample of women (i.e., Latinas born in Latin America and African American and White women). Ethnicity was self-reported based on options defined by the study investigators. The study was approved by the relevant institutional review boards, and all patients provided written informed consent.

Two hundred sixty-seven women consented to treatment and were randomized to one of the three treatment groups. The pharmacotherapy group ($n = 88$) received paroxetine, with a mean dose of 30 mg daily and a range of 10–50 mg (dosing protocol adjusted for response and reported adverse effects). The duration of this medication intervention was 6 months, in line with guidelines for the acute and maintenance phases of depression treatment (AHRQ, 1993). The study did not offer medication treatment after 6 months, but women could seek continued medication treatment elsewhere if desired. Paroxetine treatment was managed by primary care nurse practitioners under the supervision of a board-certified psychiatrist (Joyce Y. Chung). Eighteen (20%) patients unable to continue paroxetine were switched to bupropion therapy (mean dose = 229 mg/day, range = 100–450 mg).

Women in the CBT group ($n = 90$) received therapy from experienced psychotherapists who were previously trained in CBT. Therapists were supervised by a licensed clinical psychologist with CBT expertise who conducted weekly group supervision to ensure adherence to the treatment. The manual-guided treatment was eight weekly sessions administered in group or individual sessions (Muñoz, Aguilar-Gaxiola, & Guzman, 1986; Muñoz & Miranda, 1986). All patients in this arm were provided protocol-based CBT based on the course manual, and treatment involved homework and monitoring activities. Cognitive-behavior therapy could be extended an additional 8 weeks if the patient still met criteria for major depressive disorder and wanted additional therapy (15 [17%] received an additional course of CBT).

Therapists attempted to get each woman randomized to CBT into group care for cost-effectiveness reasons. When strong preferences or scheduling issues prevented them from joining a group, women were offered individual CBT. Of the 90 women assigned to psychotherapy, 32 (35.5%) completed a course of CBT defined as six or more CBT sessions. Fifteen of the 32 received group CBT, and 17 received individual CBT. Both groups received the same manual-guided treatment. Latinas were much more likely to receive individual CBT compared with African Americans and Whites. Eighty-three percent of Latinas who completed a course of CBT received individual CBT, compared with 14% of African Americans and Whites. Otherwise, there were no significant differences between women who received individual CBT and those who received group CBT in terms of any of the other baseline variables in Table 1.

Women in the community referral group ($n = 89$) were educated about depression and mental health treatments available in the community. Clinicians offered to make an appointment for the women at the end of the clinical interview to facilitate the referral and to speak with the mental health clinician. Approximately one quarter of the women declined referral. Referred participants were contacted by the referring clinician within 1–2 weeks of referral to encourage them to attend the community care program. All women in the WECare study were followed for 12 months regardless of whether they continued to receive study treatments.

Measures

Our primary outcome was the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960). WECare participants completed a structured version of the HDRS (Williams, 1988) by telephone at baseline, monthly for 6 months, and at Months 8, 10, and 12. Both the American Psychiatric Association (APA; 2000) and the National Institute for Health and Clinical Excellence (NICE; 2009) have recommended using HDRS cutoff values of 7, 13, 18, and 22 to classify subjects into different depression categories. Participants with HDRS scores of 7 or less are referred to as "not depressed." Cutoff values of 13, 18, 22, and ≥23 are used to classify participants into "mild," "moderate," "severe," and "very severe" depression categories, respectively. These are the names given by the APA. NICE uses different names but the same cutoff points.

Anxiety was measured at baseline, Month 6, and Month 12 using the Hamilton Anxiety Rating Scale (HAM-A; M. Hamilton, 1959), a 14-item rating scale that measures both psychic and somatic anxiety. Screening interviews assessed demographics, insurance status, income, and interest in treatment.

Sample

Demographic and clinical characteristics of the sample are presented in Table 1. The sample was made up of young minority women, the majority of whom were uninsured and living below or near the poverty level. There were no significant differences at baseline among the randomly assigned intervention groups on demographics, baseline depression, baseline anxiety, current diagnoses from the CIDI, and interest in treatment. Women randomly assigned to medications reported somewhat higher levels of depressive symptoms at baseline than did the other two groups, a difference that neared significance ($p = .06$). Based on diagnoses from the CIDI, about half the women were experiencing a mild to moderate episode and 47% a severe episode. Depression severity was determined based on responses to structured interview questions from the CIDI. In addition to a diagnosis of MDD, 46% of the women also had panic disorder, agoraphobia, social phobia, and/or generalized anxiety disorder. Most of the women were interested in receiving treatment.

Table 2 provides mean HDRS scores, percentage missing, and cumulative measurement dropout at each time point by treatment group. By Month 6, approximately 84% of participants had been retained in the study. By Month 12, the retention rate was 76%. The difference in dropout rates across the three treatment groups was not significant ($p = .27$).

Growth Mixture Modeling

A frequent characteristic of depression clinical trials (including the WECare study) is that outcomes over time are subject to

Table 1
WECare Variables at Baseline by Treatment Group

Variable	Total (n = 267)	Medication (n = 88)	CBT (n = 90)	TAU (n = 89)
Age in years, mean (*SD*)	29.3 (7.9)	28.7 (6.6)	29.8 (7.9)	29.5 (9.1)
Marital status, *n* (%)				
Married or living with partner	124 (46.4)	43 (48.9)	40 (44.4)	41 (46.1)
Widowed or separated/divorced	52 (19.5)	17 (19.3)	22 (24.4)	13 (14.6)
Never married	91 (34.1)	28 (31.8)	28 (31.1)	35 (39.3)
No. of children, mean (*SD*)	2.3 (1.4)	2.2 (1.2)	2.2 (1.5)	2.4 (1.6)
Ethnicity, *n* (%)				
Black	117 (43.8)	34 (38.6)	41 (45.6)	42 (47.2)
White	16 (6.0)	6 (6.8)	6 (6.7)	4 (4.5)
Latina	134 (50.2)	48 (54.6)	43 (47.8)	43 (48.3)
Insurance, *n* (%)				
Uninsured	173 (64.8)	55 (62.5)	58 (64.4)	60 (67.4)
Medical assistance	40 (15.0)	14 (15.9)	12 (13.3)	14 (15.7)
Private	54 (20.2)	19 (21.6)	20 (22.2)	15 (16.9)
Poverty, *n* (%)				
Below federal poverty	149 (60.0)	48 (57.1)	48 (56.5)	53 (60.2)
Near poor	88 (34.2)	33 (39.3)	27 (31.8)	28 (31.8)
Not impoverished	20 (7.8)	3 (3.6)	10 (11.8)	7 (8.0)
HDRS, mean (*SD*)	16.9 (5.2)	18.0 (5.1)	16.3 (5.1)	16.5 (5.2)
HAM-A, mean (*SD*)	15.2 (6.6)	16.2 (6.5)	15.1 (6.3)	14.3 (6.9)
CIDI 12-month diagnosis, *n* (%)[a]				
MDD mild, single	42 (15.9)	12 (13.8)	17 (18.9)	13 (14.9)
MDD moderate, single	62 (23.5)	14 (16.1)	21 (23.3)	27 (31.0)
MDD severe, single	104 (39.4)	44 (50.6)	31 (34.4)	29 (33.3)
MDD mild, recurrent	17 (6.4)	4 (4.6)	7 (7.8)	6 (6.9)
MDD moderate, recurrent	14 (5.3)	7 (8.1)	4 (4.4)	3 (3.5)
MDD severe, recurrent	19 (7.2)	5 (5.8)	7 (7.8)	7 (8.1)
Panic disorder	21 (8.0)	8 (9.2)	8 (8.9)	5 (5.8)
Agoraphobia without panic disorder	22 (8.3)	5 (5.8)	9 (10.0)	8 (9.2)
Social phobia	18 (6.8)	5 (5.8)	5 (5.6)	8 (9.2)
Generalized anxiety disorder	101 (38.3)	36 (41.4)	32 (35.6)	33 (37.9)
Interested in treatment, *n* (%)	232 (86.9)	82 (93.2)	75 (83.3)	75 (84.3)
Interested in medication, *n* (%)	148 (55.9)	60 (68.2)	47 (52.8)[b]	41 (46.6)[b]
Interested in group treatment, *n* (%)	179 (67.3)	60 (68.2)[b]	59 (65.6)	60 (68.2)

Note. WECare = Women Entering Care; CBT = cognitive-behavioral therapy; TAU = treatment as usual; HDRS = Hamilton Depression Rating Scale; HAM-A = Hamilton Anxiety Rating Scale; CIDI = Composite International Diagnostic Interview; MDD = major depressive disorder.
[a] Two medication participants and one TAU participant did not receive the CIDI. [b] One observation is missing.

considerable between-subjects heterogeneity due to the fact that patients often follow different trajectories over time. Some participants may see immediate gains, only to relapse at a later date, while others will improve gradually over time. Some participants will not improve at all. When comparing the effectiveness of different treatments, it is important to identify and take into account these different trajectories, because the effectiveness of an intervention may depend on the trajectory class of the participants. Despite the fact that heterogeneity of outcomes is common in depression studies, most analyses such as mixed-effects regression models (Gibbons et al., 1993) assume that all individuals are drawn from a single population with common population parameters (B. O. Muthén, 2004). That is, they assume that all individual trajectories vary around a single mean trajectory. This assumption goes counter to clinical observations and empirical data where variation in trajectory shapes is routinely observed. When individuals follow several different trajectory shapes, conventional repeated measures modeling may lead to a distorted assessment of treatment effects.

Growth mixture modeling (B. Muthén et al., 2002; B. Muthén & Shedden, 1999) relaxes the single population assumption to allow for parameter differences across several unobserved populations. Instead of considering individual variation around a single trajectory, a growth mixture model (GMM) allows different classes of individuals to vary around several different trajectories. In this way, growth mixture modeling may do a better job of capturing between-subjects variability because it does not require that all individuals follow the same average trajectory over time.

Once multiple trajectories have been identified, analyses can be performed to predict trajectory class as a function of other covariates. This approach is particularly useful in randomized trials because it may suggest that for some groups of individuals one treatment may be better than another treatment based on the subject's predicted trajectory. For example, if a subject's age, number of children, and ethnicity are predictive of a trajectory where outcomes are more favorable under medication rather than CBT, then one would consider treating a patient with similar characteristics with medication. On the other hand, it may be that a subject's predicted trajectory suggests that both medication and CBT are effective. In that case, either treatment can be offered. In this way, growth mixture modeling may provide insights on personalized depression treatments that are tailored based on patient characteristics as well as preferences.

Table 2
WECare Mean HDRS Scores (and Percentage Missing and Cumulative Measurement Dropout Percentage) at Each Time Point

Time	Medication (n = 88)	CBT (n = 90)	TAU (n = 89)
Baseline	17.95 (0, 0)	16.28 (0, 0)	16.48 (0, 0)
Month 1	14.00 (20, 2)	13.11 (27, 6)	12.80 (27, 4)
Month 2	10.74 (16, 5)	11.42 (27, 7)	11.30 (29, 10)
Month 3	9.60 (28, 8)	10.24 (36, 9)	13.05 (27, 11)
Month 4	9.54 (31, 9)	9.07 (38, 13)	11.81 (35, 12)
Month 5	8.62 (40, 14)	10.47 (34, 14)	11.85 (40, 13)
Month 6	9.17 (28, 18)	10.73 (33, 14)	11.92 (29, 15)
Month 8	8.07 (36, 24)	9.62 (30, 17)	11.55 (33, 18)
Month 10	9.04 (40, 27)	8.31 (31, 20)	10.92 (31, 19)
Month 12	9.71 (30, 30)	8.38 (24, 24)	10.22 (19, 19)

Note. WECare = Women Entering Care; HDRS = Hamilton Depression Rating Scale; CBT = cognitive-behavioral therapy; TAU = treatment as usual.

Growth mixture models have been used in a variety of applied settings where multiple trajectories have been shown to exist. In the area of depression research, growth mixture modeling has identified distinct trajectories in studies of psychotherapy (Cuijpers, van Lier, van Straten, & Donker, 2005; Stulz, Lutz, Leach, Lucock, & Barkham, 2007) and in randomized trials of medication versus psychotherapy (Lutz, Stulz, & Kock, 2009). Stulz, Thase, Klein, Manber, and Crits-Christoph (2010)—in a randomized trial of medication versus psychotherapy—identified four latent trajectories and found differential treatment effects within the two trajectories characterized by moderate depression severity.

Several researchers have noted that in cases where data are drawn from a single multivariate nonnormal distribution, a GMM will identify multiple classes, suggesting the presence of multiple latent groups when in fact the data are drawn from only one population (Bauer, 2007; Bauer & Curran, 2003).

Without strong theoretical justification, it is difficult to determine empirically whether the multiple classes identified by a GMM reflect the true population structure or are artifacts of the GMM attempting to approximate homogenous, nonnormally distributed data using a mixture of normal distributions. Here we made the assumption that the latent groups identified by our GMM represented interpretable and predictable subgroups that were present prior to the enrollment of treatment. Mean intervention effects within each of these classes were represented by changes in the parameters describing these trajectories (i.e., the means and variances for slopes and quadratic components). While these assumptions are largely untestable because they depend on unobserved or latent trajectory classes, we viewed the use of GMMs as useful for interpreting where heterogeneity of impact occurs and as a tool for predicting which treatment may be appropriate for a given patient.

In order to aid in the interpretation of the classes arising from a GMM as representative of underlying heterogeneity of effects, it has been suggested that substantive theory and auxiliary information be incorporated into the analyses. In addition, GMM analyses can lead to subsequent study designs that shed more light on identified classes, leading to more confirmatory analyses (B. Muthén, 2003). We consider interpretation of our latent classes in the Discussion section.

Statistical Analysis

We modeled the WECare baseline through Month 12 HDRS scores using the methods of B. Muthén et al. (2002), who describe a framework for fitting GMMs to randomized interventions. We assumed that trajectory class membership was a quality that characterized an individual before entering the trial and was not influenced by the intervention (B. O. Muthén, Brown, Leuchter, & Hunter, 2011). This is a strong assumption and only partly testable, as latent classes cannot be known prior to observing outcomes. As noted by C. H. Brown et al. (2008), this assumption is necessary in order to make a full causal interpretation of the effect of treatment on depression within trajectory class.

Within trajectory class, we modeled the Hamilton scores using random intercept and slopes with a fixed quadratic effect. Using a fixed quadratic effect was done because this was a small effect displaying minimal variability in the data. Consistent with our assumption that treatment group does not influence class membership, we constrained the intercept means and variances to be the same across treatment groups within trajectory class but allowed them to vary between trajectory classes. Residual variance was also constrained to be the same across treatment groups within class but allowed to vary across classes.

The effect of treatment on depression was modeled by estimating the Treatment × Slope interaction and the Treatment × Quadratic Slope interaction. These parameters were allowed to vary both within and between classes. Thus, we can estimate the effects of the WECare interventions within class and see if these effects differ based on trajectory class membership. As our interest was on the comparative effectiveness of medication versus CBT, we only compared these two interventions with each other within each class. However, we did include the TAU participants in our analyses in order to decrease error variance and increase the precision of our parameter estimates.

There are two components to a GMM: a model estimating outcomes conditional on latent trajectory class and a model predicting latent trajectory class as a function of covariates. For the component of the GMM for predicting trajectory class membership as a function of baseline covariates, our choice of predictors of trajectory class membership was guided by patient characteristics identified in the literature as affecting the course of depressive symptoms or its treatment. These variables were baseline HAM-A score (level of anxiety symptoms), age, and marital status (Chen, Eaton, Gallo, & Nestadt, 2000; K. E. Hamilton & Dobson, 2002; Keller, Lavori, Lewis, & Klerman, 1983; Mueller et al., 1999; Trivedi, Morris, Pan, Grannemann, & John Rush, 2005). Because the WECare study was designed to test the effectiveness of treatment for depression among ethnic minorities, we also included ethnicity in our models predicting trajectory class.

As suggested by B. Muthén et al. (2002), we began by fitting three separate GMMs for each treatment group. This is an exploratory technique for checking whether an intervention influences class membership. In particular, fitting a GMM using data from the TAU group alone provides evidence regarding the number and shape of trajectories in the absence of treatment. If an intervention does not influence class membership—as we assumed in our analysis—then when fitting models separately by treatment group, we would expect to obtain results similar to those obtained when combining all participants into the same

analysis. We started by fitting models with two trajectories and then increased the number of trajectories until we found the number that best fit the data. We used the Bayesian information criteria (BIC; Kass & Raftery, 1995; Schwarz, 1978) to determine the correct number of trajectories, as it has been shown to be conservative (E. C. Brown, 2003) and to perform best among other information criteria for identifying the number of classes in a GMM (Nylund, Asparoutov, & Muthén, 2007). We used Mplus 5.21 (L. Muthén & Muthén, 1998–2003) to fit our GMM.

Once a final model was found, we compared the effectiveness of medication versus CBT by calculating the differences in predicted mean HDRS scores between the two treatment groups at Months 6 and 12 within each trajectory class. We chose HDRS scores at 6 months as it measures depression symptoms immediately after the end of treatment, while HDRS scores at 12 months are a better reflection of long-term response to care. We also calculated the remission rates at Months 6 and 12 among these two treatment groups by trajectory class using the definition of remission consistent with Miranda et al. (2006). Remission was calculated using predicted individual HDRS scores and was defined as an HDRS score ≤ 7 and a 50% change in HDRS scores from baseline to Month 6 or Month 12.

Since the component of our GMM that predicts class cannot include baseline depression, as this variable is a manifest variable in the latent growth model, we also performed follow-up analyses outside of our GMM to assess how well we could predict trajectory class using only baseline-level covariates including baseline depression. Our goal was to produce an algorithm that would allow practitioners to predict trajectory class prior to beginning treatment, so that they could personalize their interventions to provide the best result for their patients. To do this, we used two approaches. The first was a logistic regression predicting trajectory membership as a function of baseline covariates. Since a GMM provides probabilities of class membership representing the degree of uncertainty for each individual, we generated 20 multiple imputations of class membership using the class probabilities. We then fit our logistic regression model separately on each of the 20 multiply-imputed data sets and combined results across the multiple imputations (Little & Rubin, 2002). The second approach used a classification tree (Breiman, Friedman, Olshen, & Stone, 1984) to classify trajectory class as a function of successive decision rules by searching through a list of baseline covariates. An indicator variable for trajectory class for the decision tree was generated using one imputation from the class probabilities. The variables used in the classification tree were baseline depression, baseline anxiety, interest in treatment, interest in group treatment, interest in medication, ethnicity, age, marital status, number of children, and the CIDI diagnoses listed in Table 1. In order to find the right-sized tree that minimized classification error, we performed cost–complexity pruning with 10-fold cross-validation (Breiman et al., 1984). Logistic regression was performed using SAS, and classification trees were performed using the package rpart (Therneau & Atkinson, 2011) in the software environment R (R Development Core Team, 2005).

Results

Identification of Latent Trajectories

Figure 1 displays the results of a two-class GMM for the WECare depression scores. This model produced a better fit to the data (BIC = 12,847) compared with a conventional one-class model (BIC = 12,865). Although the standard likelihood ratio test does not follow a chi-squared distribution in this setting, the magnitude of the test statistic and its corresponding p value, $\chi^2(27) = 168.98$, $p < 10^{-21}$, also provide strong evidence that a two-class solution is preferable to a one-class solution. Fitting models separately by treatment group also pointed toward a two-class solution. A three-class GMM did not converge—as it added a class of very small size—so we retained the two-class model.

Starting from the top of Figure 1, the first three lines consist of predicted HDRS means for the TAU, CBT, and medication interventions in the first trajectory class. This class accounts for 35% of the sample. The participants in this class are characterized by their high baseline depression scores ($M = 20.1$ vs. an overall mean of 16.9). For ease of presentation, we refer to this class of participants as the Severe class. The Severe class is also characterized by modest improvements in depression over time. For those severe-class subjects in the medication group, depression worsened after Month 6, to the point where the average HDRS score at the end of the study was higher for medication participants than for any of the other participants.

The second set of three lines in Figure 1 corresponds to the predicted HDRS means for the TAU, CBT, and medication interventions in the second trajectory class. This class accounts for the remaining 65% of the sample. Participants in the second class are characterized by their moderate depression scores ($M = 13.5$) at baseline. We refer to this trajectory class as the Moderate class. Within the Moderate class, depression scores over time for all three intervention groups follow a similar trajectory, with the medication intervention having lower mean scores at every time point after baseline.

In the component of the GMM predicting latent trajectory as a function of baseline covariates, only baseline Hamilton anxiety was a significant predictor of latent class. For each additional point on the anxiety scale, the odds of being in the Severe-class trajectory increased by 31% ($p < .001$). Ethnicity and age were not significant.

Differential Treatment Effects Within Latent Trajectories

Table 3 reports predicted HDRS means by class at Months 6 and 12. A test of the difference between the medication and CBT curves in the Severe class was significant, likelihood ratio $\chi^2(3) = 11.136$, $p = .01$. At 6 months, the estimated average HDRS score in the Severe class was 13.9 for the medication intervention and 14.9 for the CBT intervention (difference = –0.9; $p = .72$). At 12 months, the estimated means were 16.4 and 11.0, respectively (difference = 5.4; $p = .04$). The difference between Month 12 and Month 6 estimated average HDRS scores for medication participants in the Severe class was 2.5 ($p = .06$). The difference for the CBT participants was –3.9 ($p = .02$). These results suggest that while there is no difference in treatment effects between the

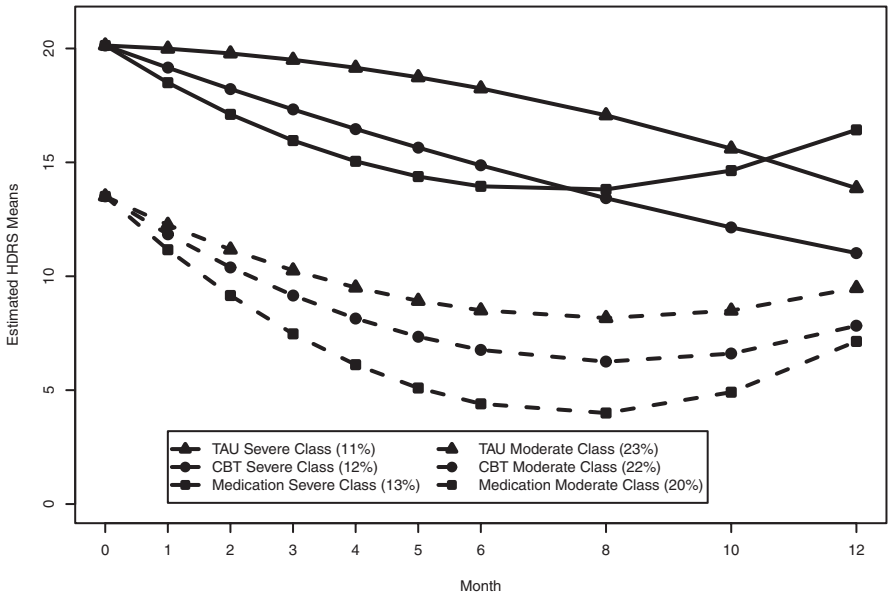

Figure 1. Two-class growth mixture model of the WECare data. The top three solid lines are the latent trajectory for the Severe class (35% of the sample), and the bottom three dotted lines are the latent trajectory for the Moderate class (65% of the sample). The difference between the medication and CBT group in the Severe class is significant at 12 months. The difference between the medication and CBT group in the Moderate class is significant at 6 months. HDRS = Hamilton Depression Rating Scale; TAU = treatment as usual; CBT = cognitive–behavioral therapy; WECare = Women Entering Care.

medication and CBT groups in the Severe class at Month 6, outcomes are significantly better for CBT participants at 1 year due to continued improvement by CBT participants after 6 months.

The difference between the medication curve and the CBT curve in the Moderate class was also significant, $\chi^2(3) = 8.684$, $p = .03$. At 6 months, the estimated means of the medication and CBT groups in the Moderate class were 4.4 and 6.8, respectively (difference = –2.4; $p = .03$). At 1 year, the estimated means were 7.1 and 7.8, respectively (difference = –0.7; $p = .61$). The difference between Month 12 and Month 6 estimated mean HDRS scores for medication participants in the Moderate class was 2.7 ($p = .003$). The difference for the CBT participants was 1.1 ($p = .21$). These results suggest that outcomes are significantly better for medication participants at 6 months but that this difference disappears by 12 months due to subsequent worsening of symptoms by medication participants after 6 months.

Remission rates in the severe trajectory class were low, due to the fact that these participants had high depression scores at baseline and were unlikely to have an HDRS score less than or equal to 7 by Months 6 or 12. In the Severe class, remission rates at 6 months for the medication and CBT groups were 0.0% and 3.5%, respectively. At 12 months they were 0.0% and 31.0%, respectively. In the moderate trajectory class, remission rates at 6 months for the medication and CBT groups were 80.0% and 54.1%, respectively. At 12 months, they were 43.6% and 41.0%, respectively.

Predicting Trajectory Membership at Baseline

Taken together, the results from our two-class GMM suggest that CBT results in better outcomes for those participants in the Severe class at Month 12, while for those in the Moderate class,

Table 3
Estimated Mean HDRS Scores by Treatment Group and Latent Trajectory Class

Time	Medication ($n = 90$)	CBT ($n = 89$)	Difference (95% CI)	Effect (95% CI)	p
		Severe latent trajectory class (35%)			
Month 6	13.9	14.9	–0.9 (–5.9, 4.1)	–0.14 (–0.89, 0.62)	.72
Month 12	16.4	11.0	5.4 (0.25, 10.6)	0.74 (0.04, 1.45)	.04
		Moderate latent trajectory class (65%)			
Month 6	4.4	6.8	–2.4 (–4.5, –0.3)	–0.49 (–0.92, –0.05)	.03
Month 12	7.1	7.8	–0.7 (–3.3, 1.9)	–0.11 (–0.50, 0.29)	.61

Note. Usual care participants were included in this model, but their results are not reported here. HDRS = Hamilton Depression Rating Scale; CBT = cognitive-behavioral therapy; CI = confidence interval.

medication results in better outcomes at Month 6 but no difference with CBT at Month 12. Our next goal was to see if we could predict trajectory class using only baseline information so that a depression intervention could be personalized for a given patient at the start of treatment. The results from the logistic regression predicting latent class are listed in Table 4. Both baseline anxiety and baseline depression were significant. For each additional point on the HAM-A, the odds of being in the severe trajectory class increased by 18%. For each additional point on the HDRS, the odds of being in the severe trajectory class increased by 22%.

Still, the logistic regression analysis reported in Table 4 conveys only that those participants with higher baseline depression and anxiety scores have greater odds of belonging in the Severe class. It does not provide a clear decision rule. As an additional analysis, we fit a classification tree using baseline data in an attempt to define a set of decision rules that predict trajectory class. This resulted in the following rule: If the baseline anxiety score is less than 19 and the baseline depression score is less than 22, then the patient belongs in the Moderate class. All other patients (those whose baseline anxiety score is 19 or greater or their baseline depression score is 22 or greater) belong in the Severe class. A graphical representation of this decision rule is displayed in Figure 2. This decision rule classified the WECare participants with 83% accuracy. It worked particularly well for the Moderate-class participants, resulting in 89% sensitivity.

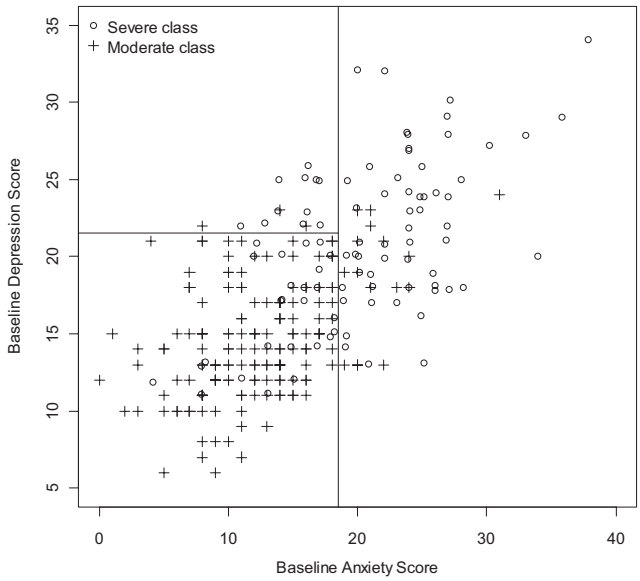

Figure 2. Graphical representation of a classification tree for predicting trajectory class. A participant with a baseline anxiety score less than 19 and a baseline depression score less than 22 is classified as being in the Moderate class. All others are considered to be in the Severe class. This rule has 83% accuracy and classifies those in the Moderate class with 89% sensitivity.

Discussion

We identified two latent trajectories of depression outcomes over the 12 months of the WECare study. Most important, the effects of antidepressant medication and CBT over a 1-year period differed based on trajectory class. Within the severe trajectory class, CBT appears to have more lasting effects, as participants receiving CBT in this class continued to improve over the course of the study while medication participants got worse after 6 months. This worsening of symptoms among the medication participants may have been due to the ending of study treatment. The average length of treatment for the medication participants in the WECare study was 4.8 months. Although not all participants began medication treatment at the beginning of the study, by 8 months almost all participants had completed treatment.

The continued improvement of CBT participants in the Severe class suggests that CBT had provided the WECare participants with a set of adaptive coping skills that allowed them to sustain their clinical gains. Within the moderate trajectory class, the medication intervention resulted in significantly better outcomes at 6 months, but this difference disappeared by 12 months. This disappearance may also have been due to ending of the study medications.

We calculated remission rates by trajectory class and found that remission rates were much higher in the Moderate class versus the Severe class. These differences are, in a sense, an artifact of the remission criteria. Remission is defined as an HDRS score ≤ 7 and a 50% change in HDRS scores from baseline to Month 6 or Month 12. Participants in the Moderate class had a mean baseline HDRS score of 13.5. On average, if participants in the Moderate class had a change in depression score of 50%, their resulting HDRS scores would be ≤ 7 and they would meet the remission criteria. On the other hand, participants in the Severe class had a mean baseline value of 20. For the average participant in the Severe class to end up with an HDRS score ≤ 7, the person needed to achieve a 65% change in HDRS score from baseline. Thus, those in the Severe class had more difficult criteria for achieving remission.

We defined a set of decision rules for classifying patients into a trajectory class so that providers can prescribe the treatment that will work best for a given patient. Depressed patients whose baseline Hamilton anxiety score is less than 19 and their baseline HDRS score is less than 22 are classified as being in the Moderate class. All others are to be assigned to the Severe class. The APA defines an HDRS score of 23 or greater as "very severe" depression (Task Force for the Handbook of Psychiatric Measures, 2000), while the NICE (2009) defines an HDRS score of 23 or greater as "severe" depression. Matza, Morlock, Sexton, Malley, and Feltner (2010) referred to HAM-A scores between 15 and 23

Table 4
Odds of Assignment Into the Severe Latent Trajectory Class

Covariate	Odds ratio (95% CI)	*p*
Baseline depression	1.18 (1.06, 1.32)	<.01
Baseline anxiety	1.22 (1.12, 1.34)	<.01
No. stressful life events	1.09 (0.90, 1.32)	.35
African American	ref	
Latina	2.17 (0.90, 5.21)	.08
White	0.85 (0.16, 4.52)	.85
Married/partnered	0.56 (0.24, 1.30)	.17
Age	0.99 (0.94, 1.04)	.71

Note. Trajectory class uncertainty was incorporated into the analysis by multiply-imputing trajectory class using trajectory class probabilities. CI = confidence interval; ref = reference group.

as "moderate" anxiety and scores greater than 24 as "severe" anxiety. Thus, our empirically derived rules for classifying patients into the two latent trajectories based on depression and anxiety scores correspond to established cutoffs and provide some indirect evidence that our latent trajectories represent true population subgroups.

The results of the participants in the severe trajectory class are consistent with those of Hollon et al. (2005), who compared severely depressed patients who responded to 16 weeks of cognitive therapy with patients who responded to 16 weeks of antidepressant therapy. After withdrawing both groups of patients from treatment, they found that those patients who had received cognitive therapy were significantly less likely to relapse over the following year.

Interestingly, the study by Hollon et al. (2005) also included a third treatment group—medication responders who were randomly assigned to continue medication. The authors found that these patients were no more likely to relapse than the cognitive therapy patients. These results suggest that those Severe class medication participants in the WECare study may have benefited by continuing medication after 6 months. This conclusion is supported by published guidelines that suggest longer treatment durations (Qaseem, Snow, Denberg, Forciea, & Owens, 2008).

As noted above, we assumed that trajectory class membership is a quality that characterizes an individual before entering the trial and is not influenced by the intervention. This strong assumption was invoked in order to make a full causal interpretation of the effect of treatment on depression within trajectory class (C. H. Brown et al., 2008). That is, in order to interpret the difference in depression scores between the medication and CBT groups at 6 and 12 months as the change in mean depression scores if a participant switched from one treatment to another at baseline, we assumed that the distribution of treatment groups was the same within a trajectory class at baseline and we constrained treatment group intercepts and variances to be the same across treatment groups within trajectory class.

The fact that we observed growth trajectories in the TAU group that were similar to those in the medication and CBT groups provides some indirect evidence that treatment does not influence trajectory class. In addition, as suggested by a reviewer, we also performed several analyses to address how sensitive our results were to our decision to constrain intercept means and variances within trajectory class by allowing intercept means and/or variances to vary across treatment group within class. From a BIC standpoint, our original model, where intercept means and variances were constrained, far outperformed models that unconstrained intercept means, variances, or both. Treatment differences at 6 and 12 months were numerically almost the same across all these models with some previously significant p values now reaching only marginal significance, which is often the case in small-sample settings when additional parameters are added to a model with no corresponding improvement in model fit. We feel that these results do not provide evidence for changing our model and limiting our ability to make causal inferences. These results also suggest that the slightly higher baseline depression scores among the medication participants had little effect on our inferences or interpretations.

Several additional limitations to this study should be noted. First, all measures were self-report, although the study measures have been used extensively in previous depression studies. Second, nearly half of the women who screened positive for depression did not follow up with diagnostic interviews, and others claimed to no longer be depressed when interviewed later. Thus, the extent to which this sample represents depressed low-income young women is not known. In general, those women who did not complete the CIDI tended to be harder to reach, less motivated to enter the study, and more likely to have their phone disconnected than were women who did complete the CIDI. We do not have reason to believe that the effect of the treatments would be different among those women who did not complete the CIDI or that these women would be less compliant with the treatments had they entered the study. Women who screened positive for depression and did not meet the criteria for MDD on the CIDI were most likely in a bad mood at screening but not clinically depressed.

A third limitation is that although the psychotherapists were closely supervised, neither adherence to the manualized treatment nor the competence of the therapists was measured. Fourth, the impact of psychotherapy or medication apart from the enhancements necessary to bring women to treatment, including providing babysitting, transportation, and up to four educational sessions conducted by the clinical providers, cannot be determined in this study. Finally, our GMM was able to identify two distinct latent trajectories. This does not necessarily mean that there are only two latent trajectories in this population of patients. A data set with more participants could provide greater power to detect additional classes, should they exist.

Our study suggests that CBT appears to be a promising treatment for young, low-income women with depression. Young, low-income women with moderate depression and anxiety are likely to improve on either antidepressant medication or CBT, with better outcomes under medication at 6 months and no difference in outcomes between the two treatments at 12 months. Those women with either severe depression or anxiety are more likely to do better on CBT. These findings are intriguing as they suggest that a short course of CBT can offer important outcomes for low-income women with significant depressive symptomology. In an attempt to verify our latent classes, future work could stratify participants by baseline depression and anxiety and investigate the effects of treatment within these strata prospectively. It would also be important to study longer term follow-up of disadvantaged women, who often continue to experience stress, to determine the impact of short-term care on important long-term outcomes, including the mental health outcomes of their children.

References

Agency for Healthcare Research and Quality. (1993). *Depression in primary care: Vol. 2. Treatment of major depression* (Clinical Practice Guideline No. 5, AHRQ Publication No. 93–0551). Washington, DC: Author.

American Psychiatric Association. (2000). *Practice guideline for the treatment of patients with major depressive disorder* (2d ed.). Washington, DC: Author.

Andrade, L., Caraveo-Anduaga, J. J., Berglund, P., Bijl, R. V., De Graaf, R., Vollebergh, W., . . . Wittchen, H.-U. (2003). The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) surveys. *International Journal of Methods in Psychiatric Research, 12(1),* 3–21. doi:10.1002/mpr.138

Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research, 42,* 757–786. doi:10.1080/00273170701710338

Bauer, D. J., & Curran, P. J. (2003). Overextraction of latent trajectory classes: Much ado about nothing? Reply to Rindskopf (2003), Muthén (2003), and Cudeck and Henly (2003). *Psychological Methods, 8,* 384–393. doi:10.1037/1082-989X.8.3.384

Blackburn, I. M., Eunson, K. M., & Bishop, S. (1986). A two-year naturalistic follow-up of depressed patients treated with cognitive therapy, pharmacotherapy and a combination of both. *Journal of Affective Disorders, 10,* 67–75. doi:10.1016/0165-0327(86)90050-9

Bortolotti, B., Menchetti, M., Bellini, F., Montaguti, M. B., & Berardi, D. (2008). Psychological interventions for major depression in primary care: A meta-analytic review of randomized controlled trials. *General Hospital Psychiatry, 30,* 293–302. doi:10.1016/j.genhosppsych.2008.04.001

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. New York, NY: Chapman & Hall/CRC.

Brown, C. H., Wang, W., Kellam, S. G., Muthén, B. O., Petras, H., Toyinbo, P., . . . Windham, A. (2008). Methods for testing theory and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug and Alcohol Dependence, 95*(Suppl. 1), S74–S104. doi:10.1016/j.drugalcdep.2007.11.013

Brown, E. C. (2003). *Estimates of statistical power and accuracy for latent trajectory class enumeration in the growth mixture model*. Unpublished manuscript, University of South Florida.

Casacalenda, N., Perry, J. C., & Looper, K. (2002). Remission in major depressive disorder: A comparison of pharmacotherapy, psychotherapy, and control conditions. *American Journal of Psychiatry, 159,* 1354–1360. doi:10.1176/appi.ajp.159.8.1354

Chen, L., Eaton, W. W., Gallo, J. J., & Nestadt, G. (2000). Understanding the heterogeneity of depression through the triad of symptoms, course and risk factors: A longitudinal, population-based study. *Journal of Affective Disorders, 59,* 1–11. doi:10.1016/S0165-0327(99)00132-9

Cuijpers, P., van Lier, P. A., van Straten, A., & Donker, M. (2005). Examining differential effects of psychological treatment of depressive disorder: An application of trajectory analyses. *Journal of Affective Disorders, 89,* 137–146. doi:10.1016/j.jad.2005.09.001

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., . . . Gallop, R. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry, 62,* 409–416. doi:10.1001/archpsyc.62.4.409

Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., & Tuason, V. B. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802–808. doi:10.1001/archpsyc.1992.01820100046009

Gibbons, R. D., Hedeker, D., Elkin, I., Waternaux, C., Kraemer, H. C., Greenhouse, J. B., . . . Watkins, J. T. (1993). Some conceptual and statistical issues in analysis of longitudinal psychiatric data: Application to the NIMH Treatment of Depression Collaborative Research Program dataset. *Archives of General Psychiatry, 50,* 739–750. doi:10.1001/archpsyc.1993.01820210073009

Hamilton, K. E., & Dobson, K. S. (2002). Cognitive therapy of depression: Pretreatment patient predictors of outcome. *Clinical Psychology Review, 22,* 875–893. doi:10.1016/S0272-7358(02)00106-X

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50–55. doi:10.1111/j.2044-8341.1959.tb00467.x

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.

Hollon, S. D., DeRubeis, R. J., Shelton, R. C., Amsterdam, J. D., Salomon, R. M., O'Reardon, J. P., . . . Gallop, R. (2005). Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. *Archives of General Psychiatry, 62,* 417–422. doi:10.1001/archpsyc.62.4.417

Institute of Medicine, Committee on Comparative Effectiveness Research Prioritization. (2009). *Initial national priorities for comparative effectiveness research*. Washington, DC: National Academies Press.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.

Keller, M. B., Lavori, P. W., Lewis, C. E., & Klerman, G. L. (1983). Predictors of relapse in major depressive disorder. *JAMA: Journal of the American Medical Association, 250,* 3299–3304. doi:10.1001/jama.1983.03340240025024

Kovacs, M., Rush, A. J., Beck, A. T., & Hollon, S. D. (1981). Depressed outpatients treated with cognitive therapy or pharmacotherapy: A one-year follow-up. *Archives of General Psychiatry, 38,* 33–39. doi:10.1001/archpsyc.1981.01780260035003

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.

Lutz, W., Stulz, N., & Kock, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders, 118,* 60–68. doi:10.1016/j.jad.2009.01.019

Matza, L. S., Morlock, R., Sexton, C., Malley, K., & Feltner, D. (2010). Identifying HAM-A cutoffs for mild, moderate, and severe generalized anxiety disorder. *International Journal of Methods in Psychiatric Research, 19,* 223–232. doi:10.1002/mpr.323

Miller, I. W., Norman, W. H., & Keitner, G. I. (1989). Cognitive-behavioral treatment of depressed inpatients: Six- and twelve-month follow-up. *American Journal of Psychiatry, 146,* 1274–1279.

Miranda, J., Chung, J. Y., Green, B. L., Krupnick, J., Siddique, J., Revicki, D. A., & Belin, T. (2003). Treating depression in predominantly low-income young minority women: A randomized controlled trial. *JAMA: Journal of the American Medical Association, 290,* 57–65. doi:10.1001/jama.290.1.57

Miranda, J., Green, B. L., Krupnick, J. L., Chung, J., Siddique, J., Belin, T., & Revicki, D. (2006). One-year outcomes of a randomized clinical trial treating depression in low-income minority women. *Journal of Consulting and Clinical Psychology, 74,* 99–111. doi:10.1037/0022-006X.74.1.99

Mueller, T. I., Leon, A. C., Keller, M. B., Solomon, D. A., Endicott, J., Coryell, W., . . . Maser, J. D. (1999). Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *American Journal of Psychiatry, 156,* 1000–1006.

Muñoz, R. F., Aguilar-Gaxiola, S., & Guzman, J. (1986). *Manual de terapia de grupo para el tratamiento cognitivo-conductal de depresion* [Manual for group therapy for the cognitive-behavioral treatment of depression]. San Francisco, CA: San Francisco General Hospital, Depression Clinic.

Muñoz, R. F., & Miranda, J. (1986). *Group therapy for cognitive-behavioral treatment of depression*. San Francisco, CA: San Francisco General Hospital, Depression Clinic.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods, 8,* 369–377. doi:10.1037/1082-989X.8.3.369

Muthén, B. O. (2004). *Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data*. Thousand Oaks, CA: Sage.

Muthén, B. O., Brown, C. H., Leuchter, A., & Hunter, A. (2011). *General approaches to analysis of course: Applying growth mixture modeling to randomized trials of depression medication*. Oxford, England: Oxford University Press.

Muthén, B., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., . . . Liao, J. (2002). General growth mixture modeling for randomized

preventive interventions. *Biostatistics, 3,* 459–475. doi:10.1093/biostatistics/3.4.459

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55,* 463–469. doi:10.1111/j.0006-341X.1999.00463.x

Muthén, L., & Muthén, B. (1998–2003). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

National Institute for Health and Clinical Excellence. (2009). *Depression: The treatment and management of depression in adults*. London, England: Author.

Nylund, K. L., Asparoutov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535–569. doi:10.1080/10705510701575396

Olfson, M., & Klerman, G. L. (1993). Trends in the prescription of antidepressants by office-based psychiatrists. *American Journal of Psychiatry, 150,* 571–577.

Qaseem, A., Snow, V., Denberg, T. D., Forciea, M. A., & Owens, D. K. (2008). Using second-generation antidepressants to treat depressive disorders: A clinical practice guideline from the American College of Physicians. *Annals of Internal Medicine, 149,* 725–733.

R Development Core Team. (2005). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464. doi:10.1214/aos/1176344136

Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., . . . Parloff, M. B. (1992). Course of depressive symptoms over follow-up: Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry, 49,* 782–787. doi:10.1001/archpsyc.1992.01820100026006

Simons, A. D., Murphy, G. E., Levine, J. L., & Wetzel, R. D. (1986). Cognitive therapy and pharmacotherapy for depression: Sustained improvement over one year. *Archives of General Psychiatry, 43,* 43–48. doi:10.1001/archpsyc.1986.01800010045006

Sox, H. C., & Greenfield, S. (2009). Comparative effectiveness research: A report from the Institute of Medicine. *Annals of Internal Medicine, 151,* 203–205.

Spitzer, R. L., Williams, J. B., Kroenke, K., Linzer, M., deGruy, F. V., III, Hahn, S. R., . . . Johnson, J. G. (1994). Utility of a new procedure for diagnosing mental disorders in primary care: The PRIME-MD 1000 study. *JAMA: Journal of the American Medical Association, 272,* 1749–1756. doi:10.1001/jama.1994.03520220043029

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology, 75,* 864–874. doi:10.1037/0022-006X.75.6.864

Stulz, N., Thase, M. E., Klein, D. N., Manber, R., & Crits-Christoph, P. (2010). Differential effects of treatments for chronic depression: A latent growth model reanalysis. *Journal of Consulting and Clinical Psychology, 78,* 409–419. doi:10.1037/a0019267

Task Force for the Handbook of Psychiatric Measures. (2000). *Handbook of psychiatric measures* (1st ed.). Washington, DC: American Psychiatric Association.

Thase, M. E., & Kupfer, D. J. (1996). Recent developments in the pharmacotherapy of mood disorders. *Journal of Consulting and Clinical Psychology, 64,* 646–659. doi:10.1037/0022-006X.64.4.646

Therneau, T., & Atkinson, B. (2011). Package "rpart." Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Trivedi, M. H., Morris, D. W., Pan, J. Y., Grannemann, B. D., & John Rush, A. (2005). What moderator characteristics are associated with better prognosis for depression? *Neuropsychiatric Disease and Treatment, 1*(1), 51–57. doi:10.2147/nedt.1.1.51.52298

U.S. Department of Health and Human Services. (2001). *Mental health: Culture, race, and ethnicity: A supplement to mental health: A report of the surgeon general*. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Office of the Surgeon General.

Wang, P. S., Demler, O., & Kessler, R. C. (2002). Adequacy of treatment for serious mental illness in the United States. *American Journal of Public Health, 92*(1), 92–98. doi:10.2105/AJPH.92.1.92

Wang, P. S., Ulbricht, C. M., & Schoenbaum, M. (2009). Improving mental health treatments through comparative effectiveness research. *Health Affairs (Millwood), 28,* 783–791. doi:10.1377/hlthaff.28.3.783

Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry, 45,* 742–747. doi:10.1001/archpsyc.1988.01800320058007

World Health Organization. (1997). *Composite International Diagnostic Interview (CIDI): Core Version 2.1: Interviewer's manual*. Geneva, Switzerland: Author.

(*Appendix follows*)

Appendix

Participant Flow in the WECare Study

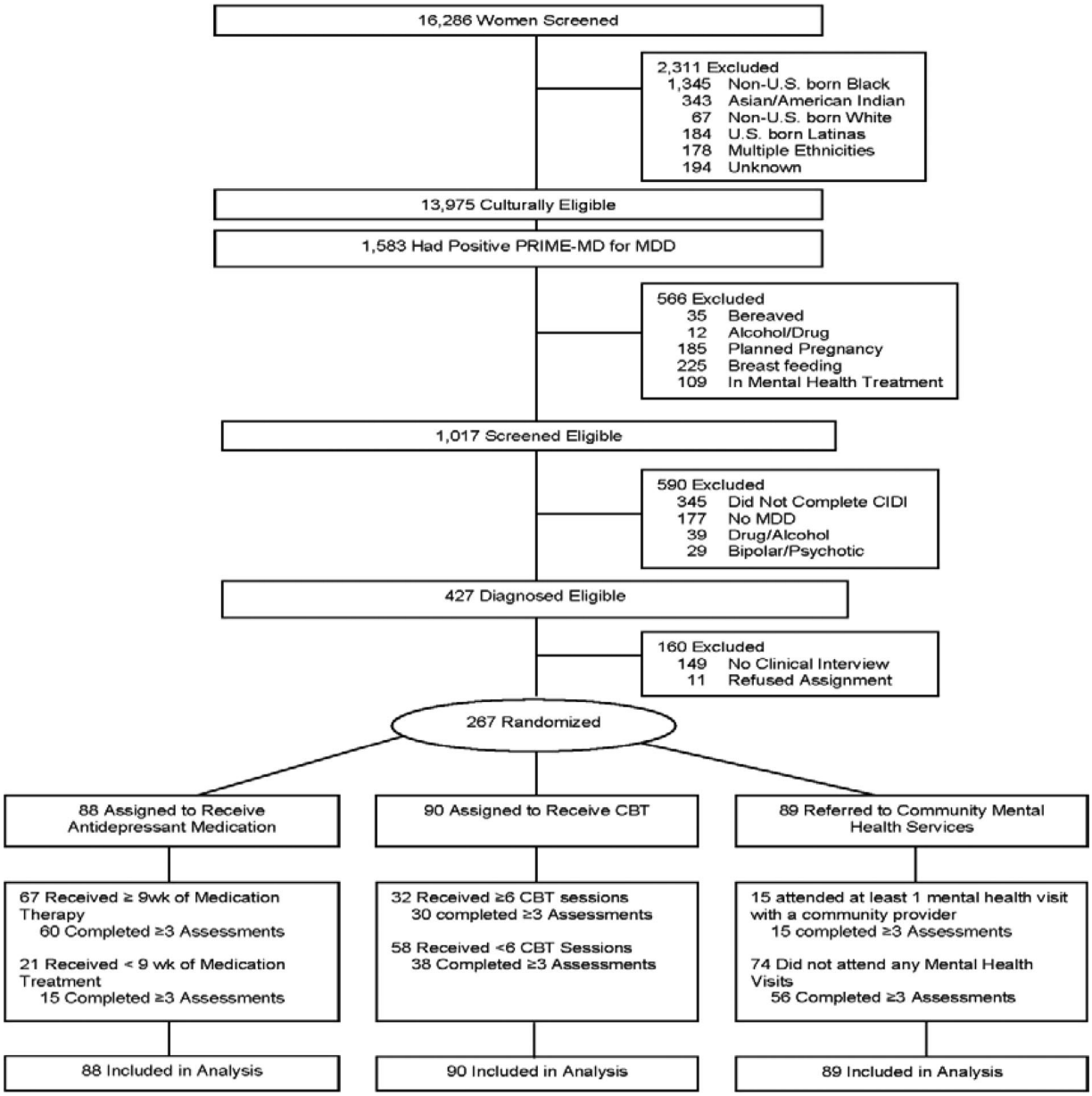

Figure A1. PRIME-MD = Primary Care Evaluation of Mental Disorders; MDD = major depressive disorder; CIDI = Composite International Diagnostic Interview; CBT = cognitive-behavioral therapy. Adapted from "Treating Depression in Predominantly Low-Income Young Minority Women: A Randomized Controlled Trial," by J. Miranda, J. Y. Chung, B. L. Green, J. Krupnick, J. Siddique, D. A. Revicki, and T. Belin, 2003, *JAMA: Journal of the American Medical Association, 290*, p. 59, Figure 1. Copyright 2003 by the American Medical Association. Adapted with permission.

Received October 20, 2011
Revision received August 20, 2012
Accepted August 23, 2012 ■