ORIGINAL ARTICLE

Trajectories of Depression Severity in Clinical Trials of Duloxetine

Insights Into Antidepressant and Placebo Responses

Ralitza Gueorguieva, PhD; Craig Mallinckrodt, PhD; John H. Krystal, MD

Context: The high percentage of failed clinical trials in depression may be due to high placebo response rates and the failure of standard statistical approaches to capture heterogeneity in treatment response.

Objective: To assess whether growth mixture modeling can provide insights into antidepressant and placebo responses in clinical trials of patients with major depression.

Design: We reanalyzed clinical trials of duloxetine to identify distinct trajectories of Hamilton Scale for Depression (HAM-D) scores during treatment. We analyzed the trajectories in the entire sample and then separately in all active arms and in all placebo arms. Effects of duloxetine hydrochloride, selective serotonin reuptake inhibitor (SSRI), and covariates on the probability of following a particular trajectory were assessed. Outcomes in different trajectories were compared using mixed-effects models.

Setting: Seven randomized double-blind clinical trials of duloxetine vs placebo and comparator SSRI.

Patients: A total of 2515 patients with major depression.

Interventions: Duloxetine and comparator SSRI.

Main Outcome Measure: Total score on the HAM-D.

Results: In the entire sample and in the antidepressant-treated subsample, we identified trajectories of responders (76.3% of the sample) and nonresponders (23.7% of the sample). However, placebo-treated patients were characterized by a single response trajectory. Duloxetine and SSRI did not differ in efficacy, and compared with placebo they significantly decreased the odds of following the nonresponder trajectory. Antidepressant responders had significantly better HAM-D scores over time than placebo-treated patients, but antidepressant nonresponders had significantly worse HAM-D scores over time than the placebo-treated patients.

Conclusions: Most patients treated with serotonergic antidepressants showed a clinical trajectory over time that is superior to that of placebo-treated patients. However, some patients receiving these medications did more poorly than patients receiving placebo. These data highlight the importance of ongoing monitoring of medication risks and benefits during serotonergic antidepressant treatment. They should further stimulate the search for biomarkers or other predictors of responder status in guiding antidepressant treatment.

Trial Registration: clinicaltrials.gov Identifier: NCT00073411

Arch Gen Psychiatry. 2011;68(12):1227-1237

Author Affiliations: Division of Biostatistics, School of Public Health (Dr Gueorguieva), and Department of Psychiatry (Dr Krystal), Yale University School of Medicine, New Haven, and Veterans Affairs National Center for Posttraumatic Stress Disorder, Veterans Affairs Connecticut Healthcare System, West Haven (Dr Krystal), Connecticut; and Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana (Dr Mallinckrodt).

THE HIGH DEGREE OF HETEROgeneity in the therapeutic response to antidepressant medications among patients with major depressive disorder constitutes an important public health problem, with significant implications for the development of medications for the treatment of depression. The presence of distinct patterns of response to antidepressant medications might help to explain why medications that were prescribed to more than 10% of the American population in 2005 are associated with modest effect sizes,[1] with only 51% of studies in the Food and Drug Administration antidepressant database showing positive results.[2,3] Further heterogeneity in clinical response in the real world may be greater than that of patients in clinical trials, as many common features of clinical complexity lead patients to be excluded from clinical trials.[4] This view is consistent with the observation that antidepressant medications, commonly prescribed and effective treatments for suicide risk among many adults,[5] increase the risk of suicide in some patients.[6]

Similarly, heterogeneous responses to placebo have plagued antidepressant drug development. High rates of placebo response[7] may contribute to the observa-

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

Table 1. Duloxetine Hydrochloride Trials for Patients With Major Depression

Protocol[a] and Trial	Total Sample Size	Treatment Arms and Sample Size per Arm			
HMAQ		**Duloxetine (20-60 mg Twice daily)**		**Fluoxetine hydrochloride (20 mg/d)**	**Placebo**
Part A	**173**	70		33	70
Part B	**193**	82		37	74
HMAT		**Duloxetine (20 mg Twice Daily)**	**Duloxetine (40 mg Twice Daily)**	**Paroxetine hydrochloride (20 mg/d)**	**Placebo**
Part A	**354**	91	84	89	90
Part B	**353**	86	91	87	89
HMAY		**Duloxetine (40 mg Twice Daily)**	**Duloxetine (60 mg Twice Daily)**	**Paroxetine hydrochloride (20 mg/d)**	**Placebo**
Part A	**367**	95	93	86	93
Part B	**392**	93	103	97	99
HMCR		**Duloxetine (60 mg Twice Daily)**		**Escitalopram oxalate (10 mg/d)**	**Placebo**
. . .	**683**	273		273	137

[a]The protocol names are an internal system that the company used to denote the study.

tion that approximately half of the studies of approved antidepressant medications produce negative or questionable results.[3] These challenges motivate research to identify distinct trajectories of clinical response to antidepressant medications or placebo on an empirical basis.[8,9] The identification of distinct trajectories of response might provide a foundation for the development of biomarkers or other predictors of treatment response, as well as refinements in the design of antidepressant clinical trials.

Trajectory-based methods address limitations of other approaches of longitudinal analysis of heterogeneous samples. For example, end-point analyses with last observation carried forward are often used, despite serious problems with bias and loss of temporal information.[10] Mixed-effects models[11] use all available data on patients, reduce bias, and improve signal detection over end-point analyses.[10,12,13] However, they assume the same mean trajectory over time for all patients within the same treatment group. Therefore, they do not account for the fact that treatment responders and nonresponders have distinct patterns of response. In contrast, trajectory-based models (also known as latent class models[14] and growth mixture models[15-17]) allow identification of distinct classes of developmental trajectories and assessment of the effect of treatments on trajectory membership. These methods have been applied to treatment research studies in the areas of alcoholism research[18,19] and depression.[20-22]

Duloxetine hydrochloride is a Food and Drug Administration–approved antidepressant medication with demonstrated efficacy for major depressive disorder.[23-25] Pooled analyses of duloxetine studies[11,25,26] have addressed important questions that required larger sample sizes than the typical clinical trial. Herein, we apply trajectory-based methods to the pooled data to explore whether they identify similar or different trajectory classes with active and placebo treatment, whether they provide new insights into the nature of antidepressant and placebo responses, and whether trajectory-based analyses improve signal detection over traditional analytic methods.

METHODS

SAMPLE

We analyzed all treatment arms from 7 randomized multicenter double-blind active and placebo comparator-controlled clinical trials of duloxetine for major depressive disorder. Earlier phase II trials conducted in the dose range of 5 to 20 mg/d that were not included in the integrated summary of efficacy were excluded. **Table 1** lists the trials, protocols, arms, and sample size per arm. Four different protocols were used for these studies (HMAQ, HMAT, HMAY, and HMCR). Parts A and B reflect trials run in parallel following the same protocol. Pooling of data from these trials was anticipated during study design. All trials incorporated double-blind variable-duration placebo lead-in periods to blind patients and investigators to the start of active therapy. Safety and efficacy results from these studies have been published previously as individual study findings[23,24,27,28] and summarized as pooled analyses of safety[25] and efficacy.[13]

For this analysis, the following 3 levels of the drug factor were used: duloxetine (all duloxetine doses), SSRI (fluoxetine hydrochloride, paroxetine hydrochloride, and escitalopram oxalate), and placebo (all placebo groups). Because protocols for dose adjustment varied across trials, we could perform dose analyses on only 2 of the protocols (HMAT and HMAY). These secondary analyses revealed no significant dose effects, and the results are not presented herein.

STATISTICAL ANALYSIS

The outcome variable was total score on the 17-item Hamilton Scale for Depression (HAM-D).[29] We used growth mixture modeling and a commercially available computer program (MPlus; Muthen and Muthen, Los Angeles, California[15-17]) to identify distinct trajectories of HAM-D scores during treatment. We first fitted models to the entire sample (duloxetine, SSRI, and placebo arms combined) and then fitted separate models to the active arms and to the placebo arms. The latter analyses were used to evaluate whether different classes would emerge for patients in the active arms and in the placebo arms. We considered linear, quadratic, and cubic trends over time, with between 1 and 4 trajectory classes. We also considered piecewise models with a change point at 2 weeks, linear change before week 2, and quadratic change after week 2.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

Table 2. Results From Model Selection for the Entire Sample

Growth Mixture Model	Likelihood	Schwartz-Bayesian Information Criterion	Entropy	Lo-Mendell-Rubin Likelihood Ratio Test: 2log Likelihood	Lo-Mendell-Rubin Likelihood Ratio Test: *P* Value	Proportion of Individuals in Class: 1	2	3	4
Linear									
1.1	−43685.38	87480.38	. . .	. . .	. . .	1.00	. . .	. . .	. . .
1.2	−43628.56	87390.24	0.56	113.63	<.001	0.78	0.22	. . .	. . .
1.3	−43607.94	87372.47	0.63	41.26	.001	0.61	0.33	0.05	. . .
1.4	−43599.64	87379.38	0.70	16.59	.08	0.61	0.33	0.05	0.004
Quadratic									
2.1	−42878.88	85898.71	. . .	. . .	. . .	1.00	. . .	. . .	. . .
2.2	−42782.65	85737.56	0.62	192.47	<.001	0.81	0.19	. . .	. . .
2.3	−42728.49	85660.56	0.68	108.32	.02	0.64	0.31	0.04	. . .
2.4	−42704.92	85644.74	0.75	47.15	.03	0.69	0.28	0.03	0.004
Cubic									
3.1	−42722.47	85593.70	. . .	. . .	. . .	1.00	. . .	. . .	. . .
3.2	−42608.06	85404.04	0.62	228.82	<.001	0.76	0.23	. . .	. . .
3.3	−42520.39	85267.85	0.72	175.34	<.001	0.74	0.23	0.03	. . .
3.4	−42440.45	85147.12	0.75	159.87	.16	0.68	0.27	0.03	0.02
Piecewise									
4.1	−42030.44	84287.95	. . .	. . .	. . .	1.00	. . .	. . .	. . .
4.2	−41931.03	84136.11	0.62	198.82	.003	0.76	0.24	. . .	. . .
4.3	−41868.67	84058.37	0.72	124.71	.31	0.69	0.29	0.02	. . .
4.4	−41819.29	84006.60	0.72	98.76	.17	0.69	0.23	0.04	0.03

The selection of the best model was based on the Schwartz-Bayesian information criterion and on the Lo-Mendell-Rubin (LMR) likelihood ratio test.[30] The LMR statistic tests whether a model with 1 less class than the fitted model describes the data as accurately and is used to select the number of classes. We applied the restriction that we need to have at least 5% of the patients in a class for that class to be meaningful clinically and stable numerically. Classification accuracy was assessed using the entropy value ranging between 0 and 1, with values closer to 1 corresponding to better classification accuracy.

Once the best-fitting model for the entire sample was identified, patients were classified to the most likely trajectory class, and weighted logistic regression analysis was performed to assess the effects of treatment and baseline characteristics on membership in a particular class. Baseline characteristics included sex, atypical flag (yes or no), melancholia flag (yes or no), age, age at onset, Hamilton Scale for Anxiety (HAM-A) total score at baseline, duration of current episode (<8 weeks, 8-18 weeks, 18 weeks to 1 year, or >1 year), and number of previous episodes (0, 1-2, 3-4, ≥5, or missing). We performed backward elimination at the $P<.10$ level. The weights were the posterior probabilities of membership in the assigned class. The association of each baseline characteristic with trajectory membership was also tested one at a time using χ^2 test, *t* test, or Wilcoxon rank sum test.

Because the separate analyses of the subsample receiving active drug and the subsample receiving placebo revealed 2 trajectory classes for patients receiving active drug (responders and nonresponders) but only 1 class for patients receiving placebo, mixed-model analyses were performed to assess whether patients receiving placebo had significantly different responses from patients receiving active drug who were classified in responder and nonresponder trajectories. To distinguish these 2 trajectory classes from the clinical definitions of responders and nonresponders, we refer to them as trajectory responders and trajectory nonresponders. The response variable in the mixed model was HAM-D total score during 8 weeks. The predictor variables were trajectory class membership (placebo, trajectory responder on duloxetine or SSRI, or trajectory nonresponder on duloxetine or SSRI), time (as a categorical variable), and their interaction.

An unstructured variance-covariance matrix was used for the errors. To control for the potential confounding of baseline covariates when comparing the randomized placebo group with the nonrandomized trajectory responder and trajectory nonresponder groups, we used a propensity scoring approach.[31] We calculated predicted probability (propensity score) for each patient to be in the trajectory nonresponder class and used this probability as a covariate in the mixed model. The propensities were calculated based on the fitted logistic regression model for patients in the active arms using all available baseline covariates as predictors.

To assess the relationship between trajectory response and clinical response, we used χ^2 test and measures of agreement. Trajectory responders were patients who were classified in the responder trajectory. Clinical responders were patients with at least 50% improvement from baseline and a HAM-D total score of less than 10 using last-observation-carried-forward imputation for missing data.

Our primary analyses are valid under missing at random (MAR) assumptions. We assessed the effect of missing data on our results by performing a limited sensitivity analysis under missing not at random (MNAR) assumptions. We used the Muthén-Roy pattern mixture model that Muthén et al[32] recommended as the most appropriate and flexible model of its class. It allows the pattern of dropout to influence the outcome of growth mixture modeling by defining 2 distinct latent class variables, one related to dropout and another related to the outcome trajectories. We identified trajectory classes based on the Muthén-Roy model and repeated all remaining analyses using the corresponding trajectory class definitions.

RESULTS

According to the Schwartz-Bayesian information criterion, the best-fitting set of models was the set of piece-

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

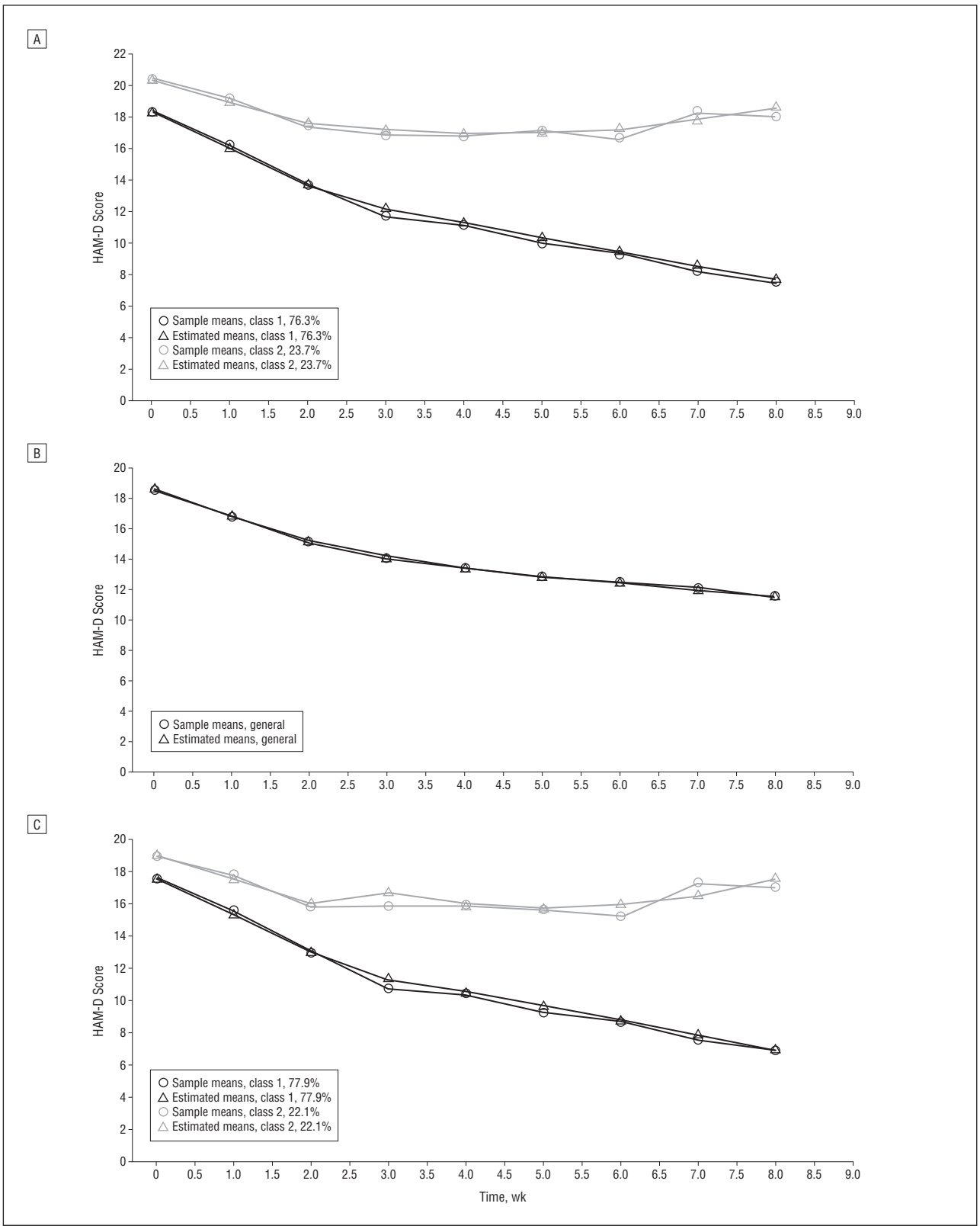

Figure 1. Sample and estimated Hamilton Scale for Depression (HAM-D) mean scores. A, For trajectory responders (class 1) and trajectory nonresponders (class 2) over time. B, For patients receiving placebo. C, For patients receiving active drug.

wise growth mixture models (**Table 2**). Among these, according to the LMR statistic, the model with 2 classes fit the data best. **Figure 1**A shows the estimated and sample means for the 2 trajectory classes over time based on all the data. The class on the bottom (class 1), with 76.3% probability of membership in this class, can be interpreted as the class of trajectory responders. The class on the top (class 2), with 23.7% probability of member-

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

Table 3. Baseline Characteristics of Trajectory Responders and Trajectory Nonresponders

Variable	Trajectory Responders (n = 2034)	Trajectory Nonresponders (n = 481)	*P* Value
Drug, No. (%)			
Duloxetine hydrochloride	978 (84.2)	183 (15.8)	<.001
SSRI	561 (79.9)	141 (20.1)	
Placebo	495 (75.9)	157 (24.1)	
Protocol, No. (%)			
HMAQ	280 (76.5)	86 (23.5)	<.001
HMAT	557 (78.8)	150 (21.2)	
HMAY	669 (88.1)	90 (11.9)	
HMCR	528 (77.3)	155 (22.7)	
Sex, No. (%)			.55
Female	1337 (80.5)	323 (19.5)	
Male	697 (81.5)	158 (18.5)	
Atypical flag, No. (%)			.67
Yes	72 (82.8)	15 (17.2)	
No	1941 (80.9)	457 (19.1)	
Melancholia flag, No. (%)			
Yes	1343 (79.5)	346 (20.5)	.02
No	683 (83.5)	135 (16.5)	
Duration of current episode, No. (%)			
<8 wk	482 (85.9)	79 (14.1)	<.001
8-18 wk	571 (85.1)	100 (14.9)	
18 wk to 1 y	489 (78.4)	135 (21.6)	
>1 y	488 (74.5)	167 (25.5)	
Previous episodes, No. (%)			
0	464 (83.8)	90 (16.2)	.06
1-2	677 (82.1)	148 (17.9)	
3-4	334 (80.5)	81 (19.5)	
≥5	390 (77.5)	113 (22.5)	
Missing	169 (77.5)	49 (22.5)	
Continuous variables, mean (SD)			
Age	42.47 (12.14)	43.38 (12.30)	.14
Age at onset	31.88 (13.87)	31.00 (13.80)	.21
HAM-A total score	15.39 (5.56)	17.71 (5.82)	<.001

Abbreviations: HAM-A, Hamilton Scale for Anxiety; SSRI, selective serotonin reuptake inhibitor.

ship in this class, can be interpreted as the class of trajectory nonresponders.

Univariate associations between classifications in the 2 trajectory classes (trajectory responder and trajectory nonresponder) and treatment, protocol, and covariates are summarized in **Table 3**. Based on logistic regression analysis with backward elimination, the following variables were significantly related to trajectory membership: drug, protocol, HAM-A total score at baseline, and duration of current episode. Compared with patients receiving duloxetine, patients receiving placebo had significantly lower odds of being in the responders trajectory (odds ratio [OR], 0.56; 95% CI, 0.42-0.73). The results for patients receiving SSRI were similar and statistically significant (OR, 0.69; 95% CI, 0.51-0.93). Patients receiving duloxetine and patients receiving SSRI did not have statistically different odds of being in the responders trajectory (OR, 1.24; 95% CI, 0.93-1.64). Higher HAM-A total score at baseline was associated with significantly lower odds of being in the responders trajectory (OR, 0.90; 95% CI, 0.88-0.92). Longer duration of current episode seemed to be associated with lower odds of being in the responders trajectory ($P=.03$), but the post hoc pairwise comparisons of longer durations with the shortest duration were not statistically significant. Compared with the other protocols, protocol HMAY was associated with significantly higher odds of being in the responders trajectory. The 2 trials using this protocol (HMAY part A and HMAY part B) were different from the other trials because they were conducted in Eastern Europe, they included 6-month extensions rather than acute phase only, and the dropout rates were 9% and 19% compared with 30% to 40% for the other trials.

Because duloxetine and SSRI were not differentiated by odds of assignment to the responders trajectory, we combined these 2 groups for the separate analyses of the active and placebo arms. Analyses of both the active and placebo subsamples showed that piecewise linear models fit the data best (**Table 4**). The active arms identified classes similar to the ones based on the entire sample. Figure 1B and C show the best-fitting solutions for the placebo and active data, respectively. Notably, the placebo arms failed to identify more than 1 trajectory class based on the LMR statistic, suggesting that there was no categorical difference in patients' responses to placebo. Rather, continuous bell-shaped distributions were sufficient to describe the between-patient heterogeneity in means, rates of change, and curvature in the response over time in the placebo arms.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

Table 4. Results From Fitting of Piecewise Linear Models to the Subset of Placebo Arms and to the Subset of Active Arms

Model	Likelihood	Schwartz-Bayesian Information Criterion	Entropy	Lo-Mendell-Rubin Likelihood Ratio Test: 2log Likelihood	Lo-Mendell-Rubin Likelihood Ratio Test: *P* Value	Proportion of Individuals in Class: 1	2	3	4
Only placebo arms									
4.1P	−11024.88	22237.68	. . .	. . .	. . .	1.00	. . .	. . .	. . .
4.2P	−10999.68	22226.16	0.94	50.40	.12	0.96	0.04	. . .	. . .
4.3P	−10986.31	22238.30	0.90	26.74	.64	0.93	0.03	0.04	. . .
4.4P	−10969.92	22244.40	0.72	32.79	.52	0.57	0.38	0.03	0.02
Only active arms									
4.1A	−30967.15	62152.67	. . .	. . .	. . .	1.00	. . .	. . .	. . .
4.2A	−30869.97	62003.49	0.67	194.36	<.001	0.77	0.23	. . .	. . .
4.3A	−30811.50	61931.74	0.77	116.93	.002	0.74	0.25	0.01	. . .
4.4A	−30771.27	61896.45	0.80	80.47	.19	0.73	0.24	0.02	0.01

To assess whether this result might be affected by potential confounding factors (eg, the number of treatment arms, percentage of patients randomized to placebo,[33] or different sample sizes for the active and placebo groups), we performed separate analyses for the 3-arm and 4-arm trials, separate analyses of trials with different placebo response rates (20%, 25%, and 40%), and separate analyses of the duloxetine and SSRI arms. The substantive conclusions were the same. Only 1 class was identified for the placebo group ($P>.1$ for all by LMR test). Separate analyses of the duloxetine and SSRI groups favored a 2-class solution over a single-class solution ($P>.03$ for duloxetine and $P>.002$ for comparator SSRI). Adding a third class did not improve the fit of these data ($P=.07$ for duloxetine and $P=.15$ for comparator SSRI), and the third class consisted of a very small percentage of patients with unstable trajectory estimation. Therefore, 2 classes were necessary to adequately describe heterogeneity in response trajectories with duloxetine and comparator SSRI.

The mixed models comparing HAM-D scores over time of trajectory responders receiving active drug, trajectory nonresponders receiving active drug, and patients receiving placebo (**Figure 2**) showed a significant interaction between trajectory class membership and time ($F_{16,1801}=73.3$, $P<.001$). Not surprisingly, trajectory responders receiving active drug showed reduced HAM-D scores compared with patients receiving placebo at week 8 ($t_{2081}=-4.08$, $P<.001$). However, pairwise comparisons at week 8 also showed significantly higher HAM-D scores for trajectory nonresponders receiving active drug compared with patients receiving placebo ($t_{2056}=7.95$, $P<.001$). When controlling for the propensity of a patient to be a trajectory responder, the differences between trajectory nonresponders receiving active drug vs patients receiving placebo and between trajectory responders receiving active drug vs patients receiving placebo at week 8 remained significant ($t_{2062}=21.23$ and $t_{2045}=-3.94$, respectively; $P<.001$ for both). Therefore, it seems that the baseline characteristics we studied were insufficient to account for the differences observed between the nonresponders receiving active drug and the patients receiving placebo.

Using the definition of clinical response as at least 50% improvement from baseline and a HAM-D total score of less than 10 and a last-observation-carried-forward imputation method for missing data, the correspondence between trajectory responders and clinical responders indicated that 480 of 481 (99.8%) trajectory nonresponders were also clinical nonresponders. However, 1318 of 2034 (64.8%) trajectory responders were classified as clinical responders. The remaining trajectory responders (716 of 2034 [35.2%]) were classified as clinical nonresponders.

Overall, dropout proportions were 23.5% of patients receiving duloxetine, 23.0% receiving SSRI, and 26.7% receiving placebo. Among trajectory responders on active drug, the dropout proportion was 18.5%, and among trajectory nonresponders receiving active drug, the dropout proportion was 24.4%. In the entire sample, the Muthén-Roy pattern mixture model under MNAR assumptions identified 2 classes of trajectory response similar to those identified under MAR assumptions (63.7% trajectory responders and 36.3% trajectory nonresponders) and 2 classes of dropout patterns (22.6% with high probability of dropout and 77.4% with low probability of dropout). More patients were classified as trajectory nonresponders in the MNAR analysis than in the MAR analysis (36.3% vs 19.1%). Univariate differences in baseline characteristics between trajectory responders and trajectory nonresponders similar to those given in Table 2 emerged, with 2 exceptions. There was no longer a significant difference in class membership by melancholia depression type ($P=.38$), and the trajectory responders identified under MNAR assumptions had older mean (SD) age at onset than trajectory nonresponders identified under MNAR assumptions (32.1 [13.7] vs 31.0 [14.1] years, $P=.04$). Backward elimination with MNAR trajectory class definitions resulted in the same final model as that reported for MAR assumptions, with similar ORs.

In the placebo group, 1 trajectory class and 2 dropout classes (23.3% with high probability of dropout and 76.7% with low probability of dropout) were identified, while in the active group 2 trajectory classes (68.0% trajectory responders and 32.0% trajectory nonresponders) and 2 dropout classes (19.0% with high probability of dropout and 81.0% with low probability of dropout) were identified. According to the MNAR analysis, the mixed models comparing HAM-D scores over time of trajectory responders receiving active drug, trajec-

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

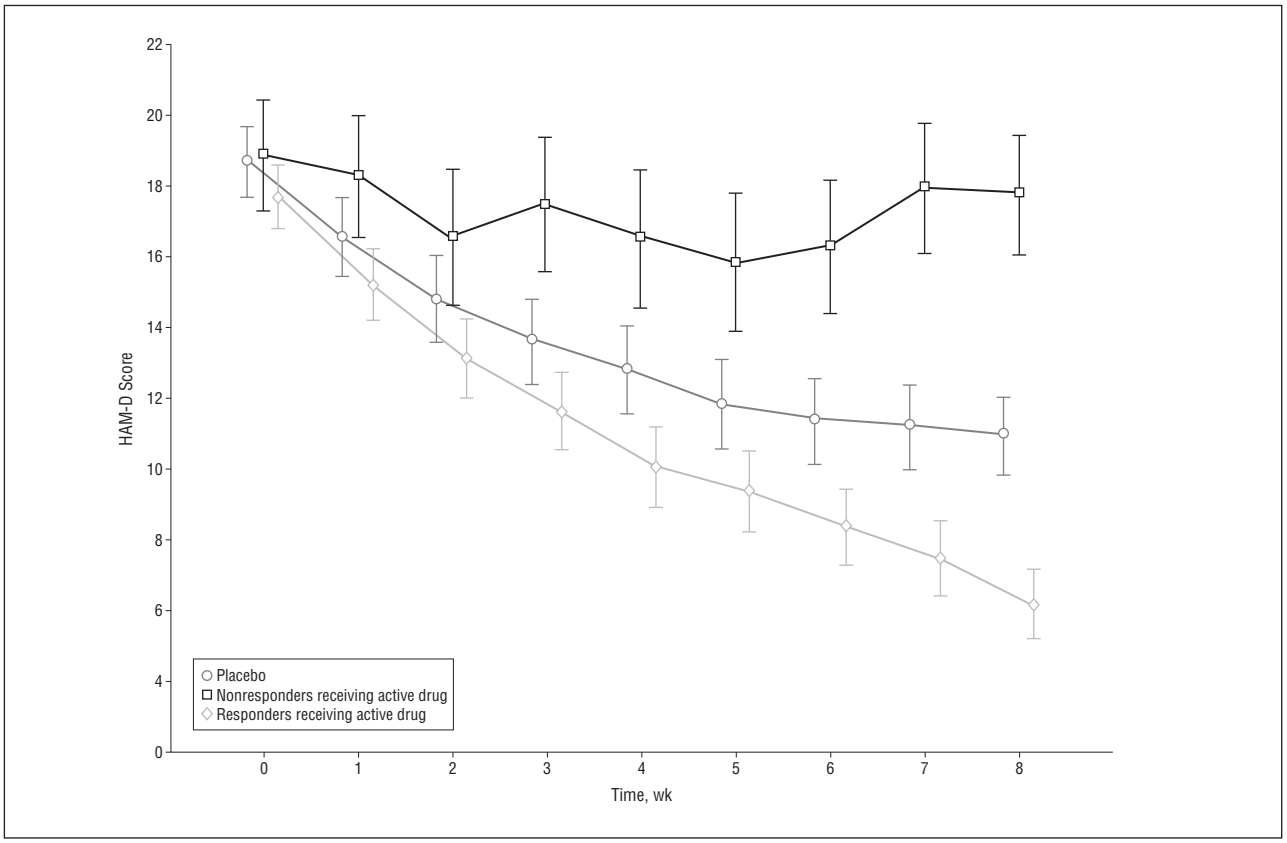

Figure 2. Least squares Hamilton Scale for Depression (HAM-D) mean (95% CI) scores for patients receiving placebo, trajectory nonresponders receiving active drug, and trajectory responders receiving active drug.

tory nonresponders receiving active drug, and patients receiving placebo showed a significant interaction between trajectory class membership and time ($F_{16,1826}$=82.6), significantly lower HAM-D scores for trajectory responders receiving active drug than for patients receiving placebo at week 8 (t_{2181}=21.72), and significantly higher HAM-D scores for trajectory nonresponders receiving active drug than for patients receiving placebo at week 8 (t_{2047}=−17.99) ($P<.001$ for all). Therefore, the MNAR sensitivity analyses confirmed our results under MAR assumptions.

COMMENT

The overall trajectory-based analyses of the treatment of 2515 patients successfully classified them into response trajectories and confirmed that duloxetine and SSRI treatment increased the likelihood that most patients treated with these medications would be classified in the responders trajectory. The magnitude of the effects seemed as large as the magnitude of the effects from end-point or mixed-model analyses.[12,13] Therefore, trajectory analyses allowed for strong signal detection, although treatment effects were not more significant than in end-point and mixed-model analyses. The added improvements of trajectory analyses were that patients were classified into response trajectories and that the trajectories were different for active drug and for placebo. As noted earlier, "responder" in this study refers to a favorable clinical trajectory rather than achievement of a priori criteria based on symptom thresholds, an approach that is commonly used in clinical trials.[34]

Separate analyses of the active and placebo groups revealed that distinct trajectories were identified in the groups treated with duloxetine or SSRI but not in the placebo group. Most patients (about three-quarters) receiving active drug were classified as trajectory responders, and fewer (about one-quarter) were classified as trajectory nonresponders. Patients receiving placebo could not be separated into distinct trajectories and on average showed gradual improvement over time. The failure to identify more than 1 trajectory classes over time in the placebo group is remarkable because growth mixture models are more prone to overestimate rather than underestimate the number of trajectory classes.[35,36] The finding of a single trajectory for patients assigned to placebo may reflect limited statistical power to resolve subtle differences in response trajectories in this group. However, this reasoning assumes that there are categorically different outcomes in each group, and this may not be the case. The present data suggest that widely divergent trajectories of individual patients treated with placebo are best explained as variations within a single class (ie, placebo response differences may be a dimensional rather than categorical characteristic).

The present findings challenge the prevailing view[37,38] that placebo response is associated predominantly with rapid and transient clinical improvement. In some investigations, researchers classified patients

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

showing this response pattern as placebo responders, although the validity of this assumption was never demonstrated. The problem with this conclusion is that growth mixture models provide valid results only under the assumption that placebo responders and nonresponders are groups differentiated by a categorical distinction. If categorically different classes do not exist, spurious latent classes are likely to be identified.[35,36] Therefore, whether one finds categorically different trajectories depends directly on whether one is willing to assume a priori that they exist. This problem is exacerbated in studies with small sample sizes. This study did not assume a priori that placebo response was constituted by multiple trajectory classes. Despite the fact that this study analyzed a much larger sample than prior studies of its kind, its analyses do not support the notion that there is a specific placebo response profile. Rather, we observed an average gradual improvement over time among patients taking placebo, with noticeable dimensional but not categorical heterogeneity. Our results are consistent with the view that placebo response is a continuous measure that is manifested to varying degrees across patients. This view is consistent with the approach by Tarpey and Petkova,[39] who used continuous rather than categorical latent variables to model patient responses to antidepressant medications.

The findings that patients receiving active drug were classified as trajectory responders and trajectory nonresponders, while there was only 1 trajectory for patients receiving placebo, may be partially explained by the effect of the active compound. Active compounds generally have the potential to produce adverse responses that could contribute to categorically different responses. For example, medications with inverted U-shaped dose-response relationships may be beneficial for patients with low baseline values but may be harmful to patients at the opposite end of the spectrum.

Another possible explanation for the single trajectory class among patients receiving placebo is that all clinical trials in this analysis had a placebo treatment phase before treatment with randomized medication. This placebo lead-in may have acted as a filter that reduced response heterogeneity during subsequent placebo treatment because it eliminated patients with immediate placebo response from the sample. Placebo lead-in is a practice that is intended to improve signal detection by selecting out patients who are likely to respond to placebo[40] and by selecting in patients whose poor placebo response portends superior medication responses.[41] However, the published data[42-46] overwhelmingly indicate that the rate of clinical response during placebo lead-in is very low and that exclusion of placebo responders does not improve signal detection in clinical trials. Therefore, it is unlikely that the presence of placebo lead-in periods in the contributing studies influenced the present findings or compromised their generalizability.

Our analyses do not support the view that antidepressant treatment response was slower and less transient than placebo response. Rather, consistent with a recent meta-analysis[47] confirming early signs of SSRI efficacy, the change in HAM-D scores as a result of antidepressant treatment among responders seemed to be faster and more sustained than the change in HAM-D scores among patients on placebo. Our analyses also support the notion that early improvement during the first 2 weeks of treatment is predictive of treatment outcome[48] and that different antidepressant medications have similar treatment response profiles.[49,50]

In the present analysis, where there are 2 trajectories for patients treated with antidepressants and 1 trajectory for patients treated with placebo, some patients would seem to be more effectively treated with placebo than with a serotonergic antidepressant. The trajectory of nonresponse to antidepressants was not consistent with the timing of a transient increase in suicidality among antidepressant-treated patients with depression.[6] Instead, the trajectories of responders and nonresponders diverged increasingly over time, suggesting that patients affected adversely by serotonin reuptake inhibitors might be better off if these medications were discontinued. At a minimum, they highlight the potential clinical importance of careful ongoing monitoring of the effect of prescribed antidepressants.

The clear separation of patients treated with antidepressants into responders and nonresponders and the observed homogeneity of treatment response on placebo, together with the significantly worse mean response of trajectory nonresponders taking active drug compared with the mean response of patients taking placebo, may help explain the failure of many clinical trials of antidepressant medication to demonstrate treatment efficacy. The main hypotheses in almost all clinical trials focus on demonstrating average treatment effects. However, when there is sizable heterogeneity in treatment response, with most patients benefiting from a treatment but with some patients demonstrating significantly worse outcomes than the average improvement while receiving placebo, it is likely that average treatment effects will be diminished and may lose statistical significance. Therefore, unless strategies are introduced to reduce this heterogeneity or study sample sizes are increased significantly to augment statistical power, it would seem that the status quo regarding failed trials of serotonin reuptake inhibitors is unlikely to change.

The fact that almost all trajectory nonresponders were clinical nonresponders but only about two-thirds of trajectory responders were clinical responders suggests that the clinical response definition is stricter than the trajectory response definition. It is possible that patients who do not meet the clinical response definition but are classified as trajectory responders may meet the clinical response definition with a longer follow-up period. This can be evaluated in follow-up analyses. Validation of our results in other studies is also necessary to assess how stable the trajectory class definitions are.

The inferences drawn from the present analyses are limited by several factors. First, we were unable to differentiate whether the poor responses to antidepressants arose from a negative effect of taking antidepressant medication or as a consequence of discontinuation of these medications. Although our limited sensitivity analysis under MNAR assumptions suggests that our results are not sensitive to the effects of missing data, further research is necessary to estimate the extent of the

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

influence of missing data due to different model assumptions for missing data.

Second, although the reported studies included only patients with unipolar major depression determined using rigorous diagnostic methods, it is impossible to rule out the presence of latent bipolar disorder in the study population. Antidepressants are reported to have reduced efficacy in treating depression among patients with bipolar disorder, and they may increase mood cycling.[51,52] Similarly, adverse effects of other forms of psychopathologic conditions, particularly personality disorders, or the social use of alcohol cannot be determined using the existing data set.

Third, the implications are limited by the brief duration (8 weeks) of the reported studies. In the present study, it was impossible to predict whether nonresponders may become responders with extended treatment or whether the negative consequences of antidepressant discontinuation in some nonresponders outweigh the apparent benefits.

Fourth, the present analysis is predicated on the assumption that heterogeneous classes of antidepressant medications exist. An alternative view is that the probability of nonresponse may vary continuously across the population, and if this is indeed the case, alternative models with latent trait rather than latent class variables might be more appropriate.[39]

Fifth, we considered only 1 type of linear parametric model. Alternative nonlinear models have been considered by other authors[53,54] to predict treatment outcome from early treatment response. It is difficult to distinguish the fit of such models from the fit of simpler polynomial models with a limited number of time points. Such models have not been used in the context of growth mixture modeling and are unavailable in standard software, but they might provide an important tool for flexible modeling of treatment response for more frequently collected treatment outcome data.

Sixth, our analyses do not assess causal treatment effects because we consider treatment a predictor of class membership rather than a predictor of growth factors within a class, as suggested by Muthén and Brown.[55] In the causal framework, latent classes are considered characteristics of the patients (eg, never responders, drug-only responders, placebo-only responders, and always responders), and a key assumption is that the "numbers of classes and the true, population proportions of patients in each class be the same across intervention conditions."[56(p596)] In our analysis, classes are considered empirically derived distinct trajectories of longitudinal response, and we have direct evidence that the number of classes and the proportion of classes vary by treatment. Although our propensity scoring approach allowed us to control for observed predictors of trajectory membership, it is possible that unmeasured confounders (eg, genotypes and environmental factors) can affect our results.

Identifying predictors of clinical trajectories might advance the personalized treatment of major depressive disorder (ie, to assist in the better matching of patients and treatments). In this regard, the present study replicated the finding from the large multicenter antidepressant trial Sequenced Treatment Alternatives to Relieve Depression (STAR*D) that high levels of anxiety predicted poorer antidepressant response.[57] Predictors of membership in both favorable and poor response trajectories might be used to inform the selection of medications for particular patients. Despite a long history of the study of clinical,[31,58] neurochemical,[59] and genetic[60] predictors of subtypes of depression and treatment response, there are still no objective bases for the personalized treatment of depression that are sufficiently explanatory and specific to guide the treatment of individual patients. Future research will be needed to determine whether trajectory-based analyses will be useful in advancing this objective.

Submitted for Publication: May 27, 2011; accepted July 8, 2011.
Correspondence: Ralitza Gueorguieva, PhD, Division of Biostatistics, School of Public Health, Yale University School of Medicine, 60 College St, Room 215B, New Haven, CT 06520-8034 (ralitza.gueorguieva@yale.edu).
Author Contributions: Dr Gueorguieva had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Financial Disclosure: Dr Krystal has been a consultant to the following companies: Aisling Capital LLC, AstraZeneca Pharmaceuticals, Biocortech, Brintnall & Nicolini Inc, Easton Associates, Gilead Sciences Inc, GlaxoSmithKline, Janssen Pharmaceuticals, Lundbeck Research USA, Medivation Inc, Merz Pharmaceuticals, MK Medical Communications, Pfizer Pharmaceuticals, F. Hoffmann-La Roche Ltd, SK Holdings Co Ltd, Sunovion Pharaceuticals Inc, Takeda Industries, Teva Pharmaceutical Industries Ltd, and Transcept Pharmaceuticals. Dr Krystal is on the scientific advisory board of Abbott Laboratories, Bristol-Myers Squibb, Eisasi Inc, Eli Lilly and Co, Forest Laboratories, Lohocla Research Corporation, Mnemosyne Pharaceuticals Inc, Naurex Inc, Pfizer Pharmaceuticals, and Shire Pharmaceuticals. Dr Krystal has the following patents and inventions: (1) patent 5 447 948 (September 5, 1995) with J. P. Seibyl and D. S. Charney on dopamine and noradrenergic reuptake inhibitors in the treatment of schizophrenia and (2) patent application PCTWO06108055A1 with G. Sanacora related to targeting the glutamatergic system for the treatment of neuropsychiatric disorders.
Funding/Support: This trial was supported by grants K05 AA 14906 and 2P50 AA 012870 from the National Institute on Alcohol Abuse and Alcoholism and the Veterans Affairs National Center for Posttraumatic Stress Disorder and by Clinical and Translational Science Award UL1 RR024139 from the National Center for Research Resources (a component of the National Institutes of Health) and the National Institutes of Health Roadmap for Medical Research.
Disclaimer: The contents herein are solely the responsibility of the authors and do not necessarily represent the official view of National Center for Research Resources or the National Institutes of Health.
Additional Contributions: Brian Pittman, MS, provided helpful comments on the manuscript.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

REFERENCES

1. Olfson M, Marcus SC. National patterns in antidepressant medication treatment. *Arch Gen Psychiatry*. 2009;66(8):848-856.
2. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med*. 2008;5(2):e45. Accessed August 24, 2011.
3. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252-260.
4. Wisniewski SR, Rush AJ, Nierenberg AA, Gaynes BN, Warden D, Luther JF, McGrath PJ, Lavori PW, Thase ME, Fava M, Trivedi MH. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. *Am J Psychiatry*. 2009;166(5):599-607.
5. Stone M, Laughren T, Jones ML, Levenson M, Holland PC, Hughes A, Hammad TA, Temple R, Rochester G. Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. *BMJ*. 2009;339:b2880 Accessed. August 24, 2011.
6. Jick H, Kaye JA, Jick SS. Antidepressants and the risk of suicidal behaviors. *JAMA*. 2004;292(3):338-343.
7. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA*. 2002;287(14):1840-1847.
8. Tarpey T, Petkova E, Odgen RT. Profiling placebo responders by self-consistent clustering of functional data. *J Am Stat Assoc*. 2003;98:850-858.
9. Gomeni R, Lavergne A, Merlo-Pich E. Modeling placebo response in depression trials using a longitudinal model with informative dropout. *Eur J Pharm Sci*. 2009;36(1):4-10.
10. Gueorguieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the *Archives of General Psychiatry*. *Arch Gen Psychiatry*. 2004;61(3):310-317.
11. Diggle PJ, Liang KY, Zeger BL. *Analysis of Longitudinal Data*. Oxford, England: Clarendon Press; 1996.
12. Mallinckrodt CH, Detke MJ, Kaiser CJ, Watkin JG, Molenberghs G, Carroll RJ. Comparing onset of antidepressant action using a repeated measures approach and a traditional assessment schedule. *Stat Med*. 2006;25(14):2384-2397.
13. Mallinckrodt CH, Prakash A, Houston JP, Swindle R, Detke MJ, Fava M. Differential antidepressant symptom efficacy: placebo-controlled comparisons of duloxetine and SSRIs (fluoxetine, paroxetine, escitalopram). *Neuropsychobiology*. 2007;56(2-3):73-85.
14. Nagin DS. Analyzing developmental trajectories: a semi-parametric, group-based approach. *Psychol Methods*. 1999;4:139-157.
15. Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*. 1999;55(2):463-469.
16. Muthén BO, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res*. 2000;24(6):882-891.
17. Muthén B, Asparouhov T. Growth mixture modeling: analysis with non-gaussian random effects. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, eds. *Longitudinal Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC Press; 2009:143-165.
18. Gueorguieva R, Wu R, Pittman B, Cramer J, Rosenheck RA, O'Malley SS, Krystal JH. New insights into the efficacy of naltrexone based on trajectory-based reanalyses of two negative clinical trials. *Biol Psychiatry*. 2007;61(11):1290-1295.
19. Gueorguieva R, Wu R, Donovan D, Rounsaville BJ, Couper D, Krystal JH, O'Malley SS. Naltrexone and combined behavioral intervention effects on trajectories of drinking in the COMBINE study. *Drug Alcohol Depend*. 2010;107(2-3):221-229.
20. Muthén B, Brown H, Leuchter A, Hunter A. General approaches to analysis of course: applying growth mixture modeling to randomized trials of depression medication. In: Shrout PE, ed. *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*. Washington, DC: American Psychiatric Publishing; 2008.
21. Hunter AM, Muthén BO, Cook IA, Leuchter AF. Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. *J Psychiatr Res*. 2010;44(2):90-98.
22. Uher R, Muthén B, Souery D, Mors O, Jaracz J, Placentino A, Petrovic A, Zobel A, Henigsberg N, Rietschel M, Aitchison KJ, Farmer A, McGuffin P. Trajectories of change in depression severity during treatment with antidepressants. *Psychol Med*. 2010;40(8):1367-1377.
23. Goldstein DJ, Mallinckrodt C, Lu Y, Demitrack MA. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *J Clin Psychiatry*. 2002;63(3):225-231.
24. Goldstein DJ, Lu Y, Detke MJ, Wiltse CG, Mallinckrodt CH, Demitrack MA. Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. *J Clin Psychopharmacol*. 2004;24(4):389-399.
25. Hudson JI, Wohlreich MM, Kajdasz DK, Mallinckrodt CH, Watkin JG, Martynov OV. Safety and tolerability of duloxetine in the treatment of major depressive disorder: analysis of pooled data from eight placebo-controlled clinical trials. *Hum Psychopharmacol*. 2005;20(5):327-341.
26. Shelton RC, Prakash A, Mallinckrodt CH, Wohlreich MM, Raskin J, Robinson MJ, Detke MJ. Patterns of depressive symptom response in duloxetine-treated outpatients with mild, moderate or more severe depression. *Int J Clin Pract*. 2007;61(8):1337-1348.
27. Detke MJ, Wiltse CG, Mallinckrodt CH, McNamara RK, Demitrack MA, Bitter I. Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Neuropsychopharmacol*. 2004;14(6):457-470.
28. Perahia DG, Wang F, Mallinckrodt CH, Walker DJ, Detke MJ. Duloxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Psychiatry*. 2006;21(6):367-378.
29. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56-62.
30. Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. *Biometrika*. 2001;88:767-778.
31. D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998;17(19):2265-2281.
32. Muthén B, Asparouhov T, Hunter A, Leuchter A. Growth modeling with non-ignorable dropout: alternative analyses of the STAR*D antidepressant trial. *Psychol Methods*. 2011;61(1):17-33.
33. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? a meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol*. 2009;19(1):34-40.
34. Tedlow J, Fava M, Uebelacker L, Nierenberg AA, Alpert JE, Rosenbaum J. Outcome definitions and predictors in depression. *Psychother Psychosom*. 1998;67(4-5):266-270.
35. Bauer DJ, Curran PJ. Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychol Methods*. 2003;8(3):338-363.
36. Tarpey T, Yun D, Petkova E. Model misspecification: finite mixture or homogeneous? *Stat Modelling*. 2008;8(2):199-218.
37. Quitkin FM, Stewart JW, McGrath PJ, Nunes E, Ocepek-Welikson K, Tricamo E, Rabkin JG, Klein DF. Further evidence that a placebo response to antidepressants can be identified. *Am J Psychiatry*. 1993;150(4):566-570.
38. Quitkin FM, McGrath PJ, Stewart JW, Taylor BP, Klein DF. Can the effects of antidepressants be observed in the first two weeks of treatment? *Neuropsychopharmacology*. 1996;15(4):390-394.
39. Tarpey T, Petkova E. Latent regression analysis. *Stat Modelling*. 2010;10(2):133-158.
40. Landin R, DeBrota DJ, DeVries TA, Potter WZ, Demitrack MA. The impact of restrictive entry criterion during the placebo lead-in period. *Biometrics*. 2000;56(1):271-278.
41. Alexopoulos GS, Kanellopoulos D, Murphy C, Gunning-Dixon F, Katz R, Heo M. Placebo response and antidepressant response. *Am J Geriatr Psychiatry*. 2007;15(2):149-158.
42. Mallinckrodt CH, Meyers AL, Prakash A, Faries DE, Detke MJ. Simple options for improving signal detection in antidepressant clinical trials. *Psychopharmacol Bull*. 2007;40(2):101-114.
43. Reimherr FW, Ward MF, Byerley WF. The introductory placebo washout: a retrospective evaluation. *Psychiatry Res*. 1989;30(2):191-199.
44. Trivedi MH, Rush H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994;11(1):33-43.
45. Faries DE, Heiligenstein JH, Tollefson GD, Potter WZ. The double-blind variable placebo lead-in period: results from two antidepressant clinical trials. *J Clin Psychopharmacol*. 2001;21(6):561-568.
46. Yang H, Cusin C, Fava M. Is there a placebo problem in antidepressant trials? *Curr Top Med Chem*. 2005;5(11):1077-1086.
47. Taylor MJ, Freemantle N, Geddes JR, Bhagwagar Z. Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. *Arch Gen Psychiatry*. 2006;63(11):1217-1223.
48. Szegedi A, Jansen WT, van Willigenburg AP, van der Meulen E, Stassen HH, Thase ME. Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients. *J Clin Psychiatry*. 2009;70(3):344-353.
49. Stassen HH, Angst J, Hell D, Scharfetter C, Szegedi A. Is there a common resilience mechanism underlying antidepressant drug response? evidence from 2848 patients. *J Clin Psychiatry*. 2007;68(8):1195-1205.

©2011 American Medical Association. All rights reserved.

Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014

50. Stassen HH, Angst J, Delini-Stula A. Fluoxetine versus moclobemide: cross-comparison between the time courses of improvement. *Pharmacopsychiatry*. 1999; 32(2):56-60.
51. Sachs GS, Nierenberg AA, Calabrese JR, Marangell LB, Wisniewski SR, Gyulai L, Friedman ES, Bowden CL, Fossey MD, Ostacher MJ, Ketter TA, Patel J, Hauser P, Rapport D, Martinez JM, Allen MH, Miklowitz DJ, Otto MW, Dennehy EB, Thase ME. Effectiveness of adjunctive antidepressant treatment for bipolar depression. *N Engl J Med*. 2007;356(17):1711-1722.
52. Truman CJ, Goldberg JF, Ghaemi SN, Baldassano CF, Wisniewski SR, Dennehy EB, Thase ME, Sachs GS. Self-reported history of manic/hypomanic switch associated with antidepressant use: data from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *J Clin Psychiatry*. 2007;68(10): 1472-1479.
53. Gomeni R, Merlo-Pich E. Bayesian modeling and ROC analysis to predict placebo responders using clinical score measured in the initial weeks of treatment in depression clinical trials. *Br J Clin Pharmacol*. 2007;63(5):595-613.
54. Merlo-Pich E, Alexander RC, Fava M, Gomeni R. A new population-enrichment strategy to improve efficiency of placebo-controlled clinical trials of antidepressant drugs. *Clin Pharmacol Ther*. 2010;88(5):634-642.
55. Muthén B, Brown HC. Estimating drug effects in the presence of placebo response: causal inference using growth mixture modeling. *Stat Med*. 2009; 28(27):3363-3385.
56. Brown CH, Wang W, Kellam SG, Muthén BO, Petras H, Toyinbo P, Poduska J, Ialongo N, Wyman PA, Chamberlain P, Sloboda Z, MacKinnon DP, Windham A; Prevention Science and Methodology Group. 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 Alcohol Depend*. 2008;95(suppl 1):S74-S104.
57. Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, Biggs MM, Zisook S, Leuchter A, Howland R, Warden D, Trivedi MH. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. *Am J Psychiatry*. 2008;165(3):342-351.
58. Carroll BJ, Feinberg M, Greden JF, Tarika J, Albala AA, Haskett RF, James NM, Kronfol Z, Lohr N, Steiner M, de Vigne JP, Young E. A specific laboratory test for the diagnosis of melancholia. Standardization, validation, and clinical utility. *Arch Gen Psychiatry*. 1981;38(1):15-22.
59. Fisar Z, Raboch J. Depression, antidepressants, and peripheral blood components. *Neuro Endocrinol Lett*. 2008;29(1):17-28.
60. Laje G, Perlis RH, Rush AJ, McMahon FJ. Pharmacogenetics studies in STAR*D: strengths, limitations, and results. *Psychiatr Serv*. 2009;60(11):1446-1457.

©2011 American Medical Association. All rights reserved.
Downloaded From: http://archpsyc.jamanetwork.com/ on 11/25/2014