Factors Affecting Antidepressant Response Trajectories: A Veterans Affairs Augmentation and Switching Treatments for Improving Depression Outcomes Trial Report

Paul B. Hicks, M.D., Ph.D., Varadan Sevilimedu, M.B.B.S., Dr.P.H., Gary R. Johnson, M.S., Ilanit R. Tal, Ph.D., Peijun Chen, M.D., M.P.H., Ph.D., Lori L. Davis, M.D., Julia E. Vertrees, Pharm.D., Sidney Zisook, M.D., Somaia Mohamed, M.D., Ph.D.

Background: In this secondary analysis of the VA Augmentation and Switching Treatments for Improving Depression Outcomes (VAST-D) study we used antidepressant response trajectories to assess the association of treatment and multiple clinical/demographic factors with the probability of response.

Methods: Using data from VAST-D, a multi-site, randomized, single-blind trial with parallel-assignment to one of three treatment interventions in 1522 Veterans whose major depressive disorder was unresponsive to at least one antidepressant trial, we evaluated response patterns using group-based trajectory modeling (GBTM). A weighted multinomial logistic regression analysis with backward elimination and additional exploratory analyses were performed to evaluate the association of multiple clinical/demographic factors with the probability of inclusion into specific trajectories. Additional exploratory analyses were used to identify factors associated with trajectory group membership that could have been missed in the primary analysis.

Results: GBTM showed the best fit for depression symptom change was comprised of six trajectories, with some trajectories demonstrating minimal improvement and others showing a high probability of remission. High baseline depression and anxiety severity scores decreased, and early improvement increased, the likelihood of inclusion into the most responsive trajectory in both the GBTM and exploratory analyses.

Conclusion: While multiple factors influence responsiveness, the probability of inclusion into a specific depression symptom trajectory is most strongly influenced by three factors: baseline depression, baseline anxiety, and the presence of early improvement.

Psych Res Clin Pract. 2023; 5:131–143; doi: 10.1176/appi.prcp.20230017

Major depressive disorder (MDD) accounts for the greatest number of disability-adjusted life years among psychiatric disorders (1). Thus, optimizing pharmacological interventions for the management of MDD is a critical goal. Attempts at characterization of antidepressant treatment response have increasingly focused on analysis of response trajectories (2, 3). Using the antidepressant agent venlafaxine XR, six response trajectory groups were observed (2). That study corroborated the frequently documented finding that over one-half of patients will have limited improvement with an antidepressant trial, and identified that high baseline depression and anxiety scores predicted being in the least responsive trajectories.

HIGHLIGHTS

- In a large study of U.S. Veterans with moderate to severe depression group-based trajectory modeling demonstrated six response trajectories as the best fit for depression symptom change over time.
- A weighted multinomial logistic regression analysis with backward elimination identified multiple factors influencing antidepressant responsiveness, but response trajectories are most strongly influenced by three factors: baseline depression, baseline anxiety, and the presence of early improvement.

One recent attempt to identify improved strategies for antidepressant use was the Veterans Affairs (VA) Augmentation and Switching Treatments for Improving Depression Outcomes (VAST-D) clinical trial, which addressed whether there was an advantage to switching antidepressants rather than augmenting with a second antidepressant or an atypical antipsychotic (4). VAST-D's large sample size provides an excellent opportunity for careful characterization of treatment response patterns.

In this secondary analysis of the VAST-D data, we used group-based trajectory modeling (GBTM) to evaluate the response trajectories for VAST-D participants. GBTM is a semiparametric technique that identifies a finite number of groups (trajectories) whose members follow similar patterns of response (5–7). Although GBTM does not make any a priori assumptions about the existence of trajectories in the population, it allows the identification of early and late responders, reduces the variability of parameter estimates, and accounts for uncertainty in individual group assignments. Using GBTM with the VAST-D data, we attempted to (1) identify unique trajectories of primary and secondary outcomes during acute phase treatment, (2) characterize the response trajectories of symptom clusters, and (3) determine whether any specific VAST-D interventions or other clinical/demographic factors would influence the likelihood of inclusion in specific trajectories, either for overall response or for symptom clusters.

METHODS

Compliance

All procedures involving human subjects/patients were approved by the VA Office of Research and Development and the VA Central Institutional Review Board, and a Certificate of Confidentiality was obtained from the National Institutes of Health. Annual reviews were conducted by the VA Central Institutional Review Board, and a Data Monitoring Committee reviewed the study biannually. Adverse events were reviewed by the VA Central Institutional Review Board and Data Monitoring Committee throughout the study. All participants provided written informed consent and privacy authorization after the procedures had been fully explained.

Study Design

VAST-D was a multisite, randomized, single-blind, parallel-assignment, next-step trial in veterans whose MDD was inadequately responsive to at least one course of antidepressant treatment with a selective serotonin reuptake inhibitor (SSRI), serotonin and norepinephrine reuptake inhibitor (SNRI), or mirtazapine that met or exceeded minimal treatment standards for dose and duration (2). Inadequate response was defined as a Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C; 8) score ≥16 (severe depression) after at least 6 weeks of treatment or a score ≥11 (moderate depression) after at least 8 weeks of treatment with the 3 most recent weeks at a stable dose. A full description of the overall design (including the CONSORT statement and flow diagram) was given in earlier manuscripts (4, 9).

Participants

Veterans Health Administration (VHA) patients with an MDD diagnosis were included in the study if they were at least 18 years old and were referred by a VHA clinician. Before enrollment, study clinicians confirmed the MDD diagnosis, and research staff reconfirmed diagnostic eligibility using criteria from the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (*DSM-IV-TR*) (10). Exclusion criteria included: pregnancy or breast-feeding; currently using contraindicated medications, including either study drug; a clear history of nonresponse or intolerance to bupropion-SR or aripiprazole; a primary diagnosis of bipolar, psychotic, obsessive-compulsive, dementia, or eating disorders; general medical conditions contraindicating the use of bupropion-SR or aripiprazole; serious, unstable medical conditions requiring acute treatment; meeting criteria for substance dependence that required inpatient detoxification; or in need of acute treatment because of suicide risk.

Interventions

This report addresses the acute phase of treatment from the VAST–D study, in which 1522 veterans with MDD were randomized to one of three treatment groups: (1) augmenting an SSRI/SNRI/mirtazapine with bupropion SR (Aug-BUP), (2) augmenting an SSRI/SNRI/mirtazapine with aripiprazole (Aug-ARI), or (3) switching to another antidepressant, that is, bupropion-SR (Switch-BUP) (4, 9). Treatments included titration (cross-titration for the switching arm)—from standard starting daily doses of either 150 mg bupropion-SR with titration up to 400 or 2 mg aripiprazole with titration up 15 mg—until depressive symptoms remitted or side effects were intolerable. Dose adjustments were guided using the Patient Health Questionnaire (PHQ-9) (11) and Frequency, Intensity and Burden of Side Effects Rating (12) obtained at each visit (baseline and at the end of weeks 1, 2, 4, 6, 8, 10, and 12).

Assessments

Baseline Measures. The baseline measures of our analysis included demographic factors (age, education, employment status, marital status, and race/ethnicity) and clinical factors or assessments (duration of index episode, presence of a substance or alcohol abuse diagnosis by the Mini-International Neuropsychiatric Interview [M.I.N.I.] (13), Adverse Childhood Experiences Survey (14), Beck Anxiety Inventory [BAI] (15), Columbia Suicide Severity Rating Scale-Lifetime Suicidal Ideation [C-SSRS] (16), 9-item adaptation of the Brief Grief Questionnaire documenting the participants' responses to the death of a close relationship [as applicable] (17), self-rated Mixed Features

Scale based on the DSM-5 (18), Cumulative Illness Rating Scale (19), PHQ-9, Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form [Q-LES-Q-SF] (20), and QIDS-C (8)). The QIDS-C evaluates the symptoms of sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Three QIDS-C symptom clusters have been characterized: core emotional cluster (energy/fatigability, concentration/decision making, loss of interest, mood, and feelings of worthlessness), sleep cluster (mid-nocturnal insomnia, sleep-onset insomnia, and early morning insomnia), and atypical cluster (psychomotor agitation, psychomotor slowing, suicidal ideation, and hypersomnia).

Outcome Measures. The primary outcome measure, QIDS-C, was collected by an independent evaluator who was blind to the treatment assignment at baseline and each visit following randomization. The PHQ-9 was collected as a secondary measure. We used standard definitions of "response" (≥50% decrease in the baseline symptom score at the end of Week 12) and "remission" (symptom score scores ≤5 on two consecutive evaluations) (2). Early improvement was defined as a ≥20% drop from baseline QIDS-C score by the end of week 2.

Statistical analysis

Trajectory Analysis. We assumed a censored normal distribution of the outcome measures (QIDS-C or PHQ-9) (21). GBTM, performed using Proc TRAJ from SAS 9.4.2 (22), uses maximum likelihood estimation to determine group sizes, the polynomial order and drop-pattern of each trajectory, and groups of individuals following similar response pathways. Groups were added to the model in a step-wise fashion, thereby assessing each group's contribution to the overall fit of the model at each step. For every subsequent addition of a group, the log Bayes factor was calculated to assess whether the addition of that group provided a better model fit. The log Bayes factor was obtained by multiplying by two the difference in Bayesian Information Criterion (subtracting a less complex model from a more complex model) for the two models under comparison. A log Bayes factor >10 was used as a benchmark to favor the more complex model. The polynomial order for each trajectory was also obtained using the log Bayes factor as a criterion for each added order. Four a priori criteria were used to assess the adequacy of the performance of the trajectory groups identified by GBTM: (1) the average estimated posterior probabilities of group membership are at least 70% (Mean Posterior Probability of Group Membership; MPP); (2) the odds of correct classification (OCC) into a group in comparison to the odds of group membership by random assignment is ≥5; (3) The differences between estimated and the actual group proportions (DEAP) for each group are expected to be <10%; and (4) the minimum group size should be ≥5% of the total population (6).

Weighted Multinomial Logistic Regression and Exploratory Analyses. To identify factors influencing assignment to trajectory groups, we performed a weighted multinomial logistic regression analysis using the posterior probability of group membership as weight. For the weighted logistic regression analysis of both the QIDS-C and PHQ-9, we implemented a Bonferroni correction for comparison of the two measures total QIDS-C and PHQ-9 such that the acceptable type I error rate in the multinomial logistic regression analysis was set to $p < 0.025$. We did not apply a Bonferroni correction for the multinomial logistic regression analyses of QIDS-C clusters because they were considered exploratory. In other exploratory analyses, we identified additional factors associated with trajectory group membership by performing either chi-square analysis (categorical data) or an analysis of variance (continuous data). For these analyses, Bonferroni corrections were not applied. All covariates included in this analysis are listed in the Baseline Measures section. This analytic approach was repeated using PHQ-9 data. In addition, the same analyses were performed on QIDS-C clusters (core emotional, sleep, and atypical). The QIDS-C Core Emotional Cluster scores are based upon a sum of QIDS-C scores for the five following items: energy/fatigability, concentration/decision making, loss of interest, mood, and feelings of worthlessness. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. Scores range from 0 to 15, with higher scores indicating greater severity of symptoms. The QIDS-C Sleep Cluster scores are based upon a sum of QIDS-C scores for the following three items: mid-nocturnal insomnia, sleep-onset insomnia, and early morning insomnia. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. Scores range from 0 to 9, with higher scores indicating greater severity of symptoms. The QIDS-C Atypical Cluster scores are based upon a sum of QIDS-C scores for the 4 following items: psychomotor agitation, psychomotor slowing, suicidal ideation, and hypersomnia. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. The total range is from 0 to 12, with higher scores indicating greater severity of symptoms.

RESULTS

Trajectory Analysis

The optimal number of trajectory groups for the VAST-D QIDS-C or PHQ-9 data is six (Supporting Information S1: Supplement B) (6). This number is based upon the criterion that the log Bayes' factor associated with the addition of a group must be >10 for the group to qualify as a significant addition to the model. The addition of a seventh group to the model also produced a Bayes' factor >10, but resulted in the proportion present in trajectory 7 being <5%. Therefore, the model with seven trajectory groups was not chosen for either analysis. The only predetermined criterion violated in the analysis using the

QIDS-C was that the difference in the actual and the estimated proportion for trajectory 3 was >10% (22.5%) (Supporting Information S1: Supplement C). In the analysis performed on the PHQ-9, we observed that there was a difference in the actual and estimated proportions for trajectory 6 (14.3%). However, the confidence intervals of the estimates of group membership probabilities were reasonably tight for QIDS-C and PHQ-9, indicating a good fit of the model. We used the log Bayes factor criteria for all combinations of quadratic and linear trajectories (5) and optimized our model to two quadratic and four linear trajectories.

Figure 1 illustrates average QIDS-C and PHQ-9 scores over the 12-week acute phase for each of the observed group trajectories. Similar patterns of response were seen for each QIDS-C cluster (Supporting Information S1: Supplement D). The QIDS-C and PHQ-9 trajectories showed similar patterns (Tables 1 and 2). Trajectories 1–3 included nearly all remitters (99.0% and 89.1% for QIDS-C and PHQ-9 scores, respectively) but only a small percentage of non-responders (11.9% and 19.6% for QIDS-C and PHQ-9 scores, respectively). In contrast, trajectories 4–6 included most of the non-responders (88.1% and 80.4% for QIDS-C and PHQ-9 scores, respectively) but only a small proportion of remitters (1.0% and 10.9% for QIDS-C and PHQ-9 scores, respectively). Patients included in trajectories 1–3 were the least likely to withdraw because of a lack of treatment response or worsening of symptoms (20.0% and 14.5% of all withdrawing for lack of treatment response; see Supporting Information S1: Supplement E).

Weighted Multinomial Logistic Regression Analysis

The odds of inclusion into specific trajectory groups versus inclusion into the least responsive group (trajectory 6) was estimated for each of the baseline measures (Tables 3–6). Elevated baseline total QIDS-C and PHQ-9 scores were more likely to be associated with less responsive trajectories. The baseline QIDS-C severity finding was also present in all QIDS-C cluster analyses. In contrast, early improvement increased inclusion into responsive trajectories, an effect that was evident for all QIDS-C clusters. Higher baseline BAI scores were more likely to be present in the least responsive trajectories for all QIDS-C clusters. Higher baseline Q-LES-Q-SF scores marginally increased inclusion into responsive trajectories for both the total QIDS-C and PHQ-9 scores, but that effect was not observed wth any of the cluster analyses. For the QIDS-C sleep cluster, the benefits of employment were strong but advanced age and decreased duration of the index episode provided only marginal benefits. For the QIDS-C atypical cluster, a modest benefit was seen for younger age or lower lifetime suicidal ideation. Greater severity of health impairment marginally increased inclusion in the least responsive trajectories. In the atypical cluster analysis, being married or cohabiting greatly increased the likelihood of inclusion in the more responsive trajectories. In these weighted multinomial logistic regression analyses, treatment allocation had no influence on trajectory inclusion for the PHQ-9, QIDS-C, or any QIDS-C cluster.

Exploratory Analysis of Clinical and Demographic Factors

Table 7 shows the statistical significance of the influence of clinical and demographic factors on inclusion into specific trajectory groups, and Supporting Information S1: Supplement F shows the actual numbers and percentages of individuals in each trajectory group. Several clinical and demographic factors increased the likelihood of inclusion into more responsive trajectories of the QIDS-C, including

FIGURE 1. Group-based trajectory model trajectories based on QIDS-C and PHQ-9 Scores of 1522 Patients from VAST-D. Group trajectories among the participants in the VAST-D study for both QIDS-C and PHQ-9 scores. Data points are the estimated scores from the model by visit for each trajectory group. PHQ-9, Patient Health Questionnaire; QIDS-C, Quick Inventory of Depressive Symptomatology-Clinician Rated; VAST-D, VA Augmentation and Switching Treatments for Improving Depression Outcomes.

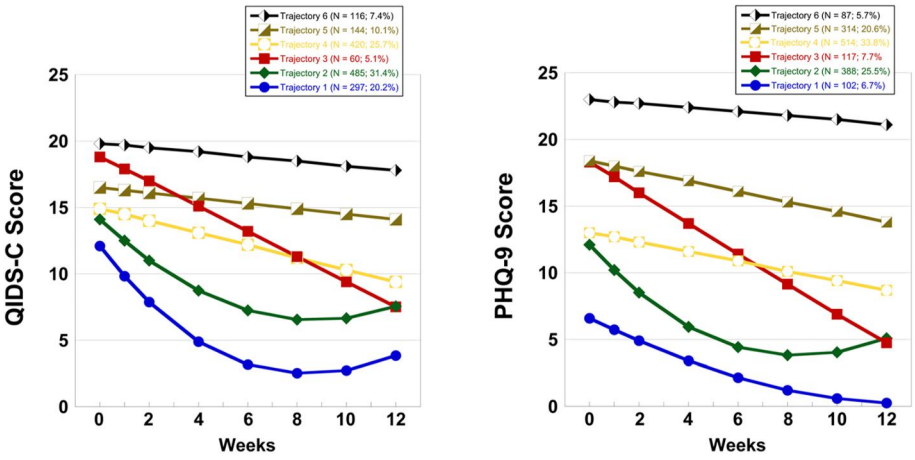

TABLE 1. Relationship between trajectory assignment, remission, and response for QIDS-C.[a]

Trajectory			Remission		Non-remission/ response		Non-response	
No.	*N*	Mean baseline QIDS-C ± SD[b]	*N*	%	*N*	%	*N*	%
1	297	14.6 ± 2.8	256	16.8	25	1.6	16	1.1
2	485	15.9 ± 2.9	133	8.7	310	20.4	42	2.8
3	60	20.9 ± 2.1	3	0.2	56	3.7	1	0.1
4	420	16.9 ± 2.7	4	0.3	194	12.7	222	14.6
5	144	18.2 ± 2.5	0	0.0	43	2.8	101	6.6
6	116	20.7 ± 2.0	0	0.0	2	0.1	114	7.5
		Totals	396	26.0	630	41.4	496	32.6

[a] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of 3 indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27.
[b] SD, Standard deviation of the mean.

TABLE 2. Relationship between trajectory assignment, remission, and response for PHQ-9.[a]

Trajectory			Remission		Non-remission/ response		Non-response	
No.	*N*	Mean baseline PHQ-9 ± SD[b]	*N*	%	*N*	%	*N*	%
1	102	9.58 ± 4.9	87	5.7	12	0.8	3	0.2
2	514	13.9 ± 4.3	243	16.0	193	12.7	78	5.1
3	117	20.9 ± 2.9	23	1.5	78	5.1	16	1.1
4	388	15.3 ± 3.5	34	2.2	231	15.2	123	8.1
5	314	19.5 ± 3.4	9	0.6	105	6.9	200	13.1
6	87	23.2 ± 2.3	0	0.0	11	0.7	76	5.0
		Totals	396	26.0	630	41.4	496	32.6

[a] PHQ-9, Patient Health Questionnaire-9. Each of nine questions is scored from 0 to 3, with 3 indicating greater severity. Possible scores on the PHQ-9 range from 0 to 27, with higher scores indicating greater degree of depression.
[b] SD, Standard deviation of the mean.

being employed, female, or Caucasian, endorsing three or fewer grief items, receiving treatment allocation Aug-ARI, experiencing a shorter index episode, having lower baseline anxiety or depression scores, fewer mixed features, a higher baseline quality of life, and lower lifetime suicidal ideation. The benefits of five of these factors were driven by contributions from all three QIDS-C clusters: being employed, having lower baseline anxiety or depression severity scores, experiencing a shorter index episode, and reporting a higher baseline quality of life. However, the influence of some factors was seen to be uniquely affected by certain QIDS-C clusters. For example, the sleep cluster influenced the race-based findings, whereas the gender-based findings had contributions from the sleep and atypical clusters. Grief endorsement effects were influenced by the sleep and core emotional clusters. The benefit of Aug-ARI treatment allocation was entirely dependent on changes in the core emotional cluster. The benefit of decreased lifetime suicidal ideation resulted from the influence on the atypical cluster, which includes the QIDS-C item on suicide ideation. Although an overall benefit was not demonstrated, a beneficial effect of having fewer mixed features was associated with outcomes of the sleep and atypical clusters.

Factors that influenced inclusion into more responsive clusters according to the PHQ-9 included being employed, being female, endorsing three or fewer grief items, demonstrating early improvement, having lower baseline anxiety or depression severity scores, a shorter index episode, the presence of fewer mixed features, and a higher baseline quality of life.

DISCUSSION

Understanding antidepressant response patterns helps shape the decision-making process in the clinical management of depressed patients. Using GBTM, six response trajectories were identified, similar to findings by other investigators (2, 3). The same patterns were found using the QIDS-C and the PHQ-9. However, the QIDS-C provides additional utility because its response clusters can be analyzed to tease out more subtle findings regarding factors influencing outcomes. The presence of these patterns across multiple studies derived from different patient populations using diverse treatment interventions suggests these response patterns could be representative of all antidepressant trials. Previous studies focused solely on the initial acute phase of treatment, but VAST-D studied participants that were ready for a "next-step" intervention. The similarity of response patterns despite the different phases of treatment suggests that antidepressant response trajectories are comparable across treatment stages.

TABLE 3. Estimates of odds ratios derived from a weighted multinomial logistic regression analysis for the total QIDS-C[a] and PHQ-9[b] scores.

		QIDS-C			PHQ-9		
Variable	Trajectory	OR[c]	95% CI	*p* value	OR	95% CI	*p* value
Early improvement[d]	1	4.80	(4.04–5.71)	<0.001	6.20	(5.10–7.55)	<0.001
	2	3.16	(2.69–3.70)	<0.001	2.48	(2.15–2.86)	<0.001
	3	1.30	(1.14–1.49)	<0.001	1.78	(1.56–2.03)	<0.001
	4	2.36	(2.02–2.74)	<0.001	3.55	(3.05–4.12)	<0.001
	5	1.89	(1.63–2.19)	<0.001	1.57	(1.39–1.78)	<0.001
QIDS-C/PHQ-9[e]	1	0.10	(0.08–0.13)	<0.001	0.07	(0.06–0.10)	<0.001
	2	0.17	(0.14–0.21)	<0.001	0.23	(0.19–0.28)	<0.001
	3	0.80	(0.65–0.98)	0.03	0.51	(0.43–0.61)	<0.001
	4	0.24	(0.19–0.31)	<0.001	0.16	(0.13–0.19)	<0.001
	5	0.35	(0.28–0.45)	<0.001	0.50	(0.43–0.59)	<0.001
Q-LES-Q-SF[f]	1	1.06	(1.03–1.10)	<0.001	1.07	(1.03–1.11)	0.001
	2	1.05	(1.02–1.07)	<0.001	1.03	(1.00–1.07	0.049
	3	1.01	(0.98–1.03)	0.60	1.03	(1.00–1.06)	0.06
	4	1.02	(1.00–1.05)	<0.001	1.05	(1.02–1.09)	0.002
	5	1.01	(0.99–1.04)	0.22	1.02	(0.99–1.05)	0.16
Treatment allocation: Aug-ARI[g] versus Switch-BUP[h]	1	1.86	(0.74–4.68)	0.17	0.96	(0.29–3.20)	0.95
	2	1.55	(0.67–3.60)	0.30	1.17	(0.45–3.04)	0.74
	3	1.85	(0.85–4.09)	0.13	1.67	(0.67–4.18)	0.27
	4	0.76	(0.34–1.70)	0.51	1.45	(0.53–3.94)	0.46
	5	1.45	(0.66–3.20)	0.34	1.32	(0.58–3.03)	0.50
Treatment allocation: Aug-BUP[i] versus Switch-BUP	1	1.11	(0.46–2.71)	0.81	0.82	(0.26–2.61)	0.73
	2	0.96	(0.43–2.13)	0.91	0.60	(0.24–1.51)	0.27
	3	0.59	(0.24–1.43)	0.23	0.66	(0.27–1.63)	0.36
	4	0.77	(0.36–1.63)	0.49	0.66	(0.25–1.74)	0.40
	5	0.78	(0.36–1.69)	0.52	0.84	(0.38–1.87)	0.67

[a] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of 3 indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27.

[b] PHQ-9, Patient Health Questionnaire-9. Each of nine questions is scored from 0 to 3, with 3 indicating greater severity. Possible scores on the PHQ-9 range from 0 to 27, with higher scores indicating greater degree of depression.

[c] OR, Odds ratio. Odds of inclusion into a specific trajectory in comparison to the odds of inclusion into trajectory 6 (least responsive trajectory).

[d] Early improvement. The presence of a ≥20% drop from the baseline QIDS-C score by the end of week 2.

[e] QIDS-C and PHQ-9, at baseline.

[f] Q-LES-Q-SF, Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form. Possible scores range from 0% to 100% of the maximum scale score of 70, with higher scores indicating greater life satisfaction and enjoyment.

[g] Aug-ARI, allocation to augmentation with aripiprazole.

[h] Switch-BUP, allocation to switching to bupropion.

[i] Aug-BUP, allocation to augmentation with bupropion.

Identifying the factors that affect antidepressant response helps inform clinical decisions. Many clinical and demographic factors in this study were found to influence antidepressant response. Weighted multinomial logistic regression analysis for the QIDS-C, QIDS-C clusters, and the PHQ-9 shows a strong role of the severity of baseline anxiety and depression, as well as early improvement in determining responsiveness. These findings are consistent with the findings of the analysis of moderators of treatment effect for the VAST-D study (23) and with a body of literature highlighting the negative influence of a higher baseline severity of depression (23–26). However, some studies have found the degree of depression severity positively correlates with response/remission rates (27, 28) or is unrelated to inclusion into specific response trajectories (3). Clinically, it makes sense that the more severely depressed patients would be the least responsive to treatment. In addition, multiple studies, in agreement with the current findings, have demonstrated the negative effects of higher baseline anxiety levels (23, 29–31). Finally, in contrast to the findings of Uher et al. (3), most studies, including the present study, have demonstrated that the lack of early improvement predicts non-response and non-remission (32–36).

In our exploratory analyses, being employed increased the likelihood of responsiveness in the QIDS-C, all QIDS-C clusters, and the PHQ-9. Other studies have also documented the positive effects of employment (23, 25, 31, 37). Although our data does not allow us to identify all related factors, it is possible that employment could be associated with unmeasured factors (e.g. the degree of life engagement or having a sense of purpose). This finding also raises the question of whether encouraging and supporting employment could be a positive therapeutic intervention. The presence of a life partner would be expected to provide stability and improve psychological health (26). It is,

TABLE 4. Estimates of odds ratios derived from a weighted multinomial logistic regression analysis for the QIDS-C[a] core emotional cluster.

Factor	Trajectory group	OR[b]	95% confidence limits	*p* value
BAI[c]				0.03
	1	0.23	0.14–0.40	
	2	0.36	0.23–0.57	
	3	0.66	0.42–1.06	
Early improvement[d]				<0.001
	1	3.38	2.95–3.87	
	2	2.13	1.90–2.39	
	3	1.61	1.45–1.78	
QIDS-C[e]				<0.001
	1	0.34	0.30–0.39	
	2	0.49	0.43–0.54	
	3	0.69	0.63–0.76	
Treatment allocation				0.21
Aug-ARI[f] versus Switch-BUP[g]	1	1.98	0.99–3.99	
	2	1.24	0.68–2.25	
	3	0.91	0.54–1.55	
Aug-BUP[h] versus Switch-BUP	1	1.18	0.59–2.36	
	2	0.96	0.54–1.71	
	3	0.88	0.53–1.47	

[a] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated Core Emotional Cluster. Possible scores are based upon a sum of QIDS-C scores for the 5 following items: energy/fatigability, concentration/decision making, loss of interest, mood, and feelings of worthlessness. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. Scores range from 0 to 15, with higher scores indicating greater severity of symptoms.
[b] OR, Odds ratio. Odds of inclusion into a specific trajectory in comparison to the odds of inclusion into trajectory 4 (least responsive trajectory).
[c] BAI, Beck Anxiety Inventory, at baseline. Possible scores range from 0 to 3 (average rating of each of the 21 items), with higher scores indicating greater anxiety.
[d] Early improvement. The presence of a ≥20% drop from the baseline QIDS-C score by the end of week 2.
[e] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated, at baseline. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of 3 indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27.
[f] Aug-ARI, allocation to augmentation with aripiprazole.
[g] Switch-BUP, allocation to switching to bupropion.
[h] Aug-BUP, allocation to augmentation with bupropion.

TABLE 5. Estimates of odds ratios derived from a weighted multinomial logistic regression analysis for the QIDS-C[a] sleep cluster.

Factor	Trajectory group	OR[b]	95% confidence limits	*p* value
Age[c]				0.035
	1	1.03	1.00–1.05	
	2	1.02	1.00–1.04	
	3	1.00	0.98–1.03	
	4	1.03	1.01–1.05	
	5	1.01	0.99–1.02	
BAI[d]				<0.001
	1	0.23	0.14–0.40	
	2	0.36	0.23–0.57	
	3	0.66	0.42–1.06	
	4	0.30	0.19–0.48	
	5	0.53	0.38–0.72	
Duration of index episode[e]				0.031
	1	1.00	1.00–1.00	
	2	1.00	0.97–1.00	
	3	1.00	1.00–1.00	
	4	1.00	1.00–1.00	
	5	1.00	1.00–1.00	
Early improvement[f]				<0.001
	1	1.39	1.26–1.54	
	2	1.38	1.26–1.52	
	3	1.14	1.03–1.27	
	4	1.01	0.92–1.12	
	5	1.17	1.08–1.25	

continued

TABLE 5, *continued*

Factor	Trajectory group	OR[b]	95% confidence limits	*p* value
Employment status				0.005
Retired versus employed[g]	1	0.28	0.14–0.58	
	2	0.49	0.25–0.97	
	3	1.00	0.45–2.19	
	4	0.31	0.16–0.60	
	5	0.61	0.36–1.03	
Unemployed versus employed	1	0.37	0.20–0.67	
	2	0.51	0.29–0.91	
	3	0.92	0.47–1.81	
	4	0.36	0.20–0.64	
	5	0.60	0.38–0.94	
QIDS-C[h]				<0.001
	1	0.77	0.71–0.84	
	2	0.88	0.82–0.95	
	3	0.94	0.86–1.02	
	4	0.77	0.72–0.84	
	5	0.86	0.81–0.91	
Treatment allocation				0.85
Aug-ARI[i] versus Switch-BUP[j]	1	1.44	0.80–2.61	
	2	1.47	0.84–2.55	
	3	1.09	0.72–1.64	
	4	1.49	0.86–2.58	
	5	1.47	0.81–2.65	
Aug-BUP[k] versus Switch-BUP	1	1.14	0.62–2.08	
	2	1.32	0.76–2.30	
	3	1.11	0.74–1.66	
	4	1.19	0.68–2.07	
	5	1.10	0.59–2.05	

[a] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated Sleep Cluster. Possible scores are based upon a sum of QIDS-C scores for the following three items: mid-nocturnal insomnia, sleep-onset insomnia, and early morning insomnia. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. Scores range from 0 to 9, with higher scores indicating greater severity of symptoms.
[b] OR, Odds ratio. Odds of inclusion into a specific trajectory in comparison to the odds of inclusion into trajectory 6 (least responsive trajectory).
[c] Age, in years, at baseline.
[d] BAI, Beck Anxiety Inventory, at baseline. Possible scores range from 0 to 3 (average rating of each of the 21 items), with higher scores indicating greater anxiety.
[e] Duration of Index Episode, duration in months of the depression episode that is currently being treated, at baseline.
[f] Early improvement. The presence of a ≥20% drop from the baseline QIDS-C score by the end of week 2.
[g] Employment status, at baseline. The employment status by the following categories: unemployed (includes disability or assistance), retired (and not working), or employed.
[h] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated, at baseline. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of 3 indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27.
[i] Aug-ARI, allocation to augmentation with aripiprazole.
[j] Switch-BUP, allocation to switching to bupropion.
[k] Aug-BUP, allocation to augmentation with bupropion.

therefore, not surprising that marital/co-habitation benefits were found, but they were associated only with the sleep and atypical clusters. Despite that, life quality at baseline had only a marginal positive influence on trajectory assignment.

By evaluating the influence of various factors on the trajectories of QIDS-C clusters, it is apparent that clusters are often differentially driven by specific demographic and clinical factors. For example, as would be predicted by the greater likelihood of response and remission from treatment allocation to Aug-ARI in the initial acute phase intervention VAST-D analysis (4), Aug-ARI treatment allocation did increase the probability of inclusion into more responsive trajectories. The present analysis demonstrates that this important effect is primarily driven through the influence of aripiprazole on the QIDS-C core emotional cluster, which is the primary focus of treatment interventions. A similar treatment allocation effect could not be detected from PHQ-9 trajectories. In contrast, the role of mixed features and gender effects do not involve the core emotional cluster; instead, they are mediated by the sleep and atypical clusters. Also, grief endorsement effects on trajectories are driven by the core emotional and sleep clusters, not the atypical cluster. The apparent benefit of Caucasian race was driven solely by a signal from the sleep cluster

TABLE 6. Estimates of odds ratios derived from a weighted multinomial logistic regression analysis for the QIDS-C[a] atypical cluster.

Factor	Trajectory group	OR[b]	95% confidence limits	*p* value
Age[c]				0.002
	1	0.93	0.90–0.97	
	2	0.97	0.94–1.00	
	3	0.97	0.94–1.00	
BAI[d]				<0.001
	1	0.31	0.14–0.67	
	2	0.73	0.40–1.34	
	3	1.07	0.62–1.84	
CIRS severity index[e]				0.04
	1	1.12	1.03–1.21	
	2	1.03	0.97–1.11	
	3	1.03	0.97–1.10	
Early improvement[f]				<0.001
	1	0.77	0.59–0.99	
	2	1.38	1.12–1.69	
	3	1.01	0.84–1.21	
Lifetime suicidal ideation[g]				0.03
	1	0.83	0.68–1.00	
	2	0.79	0.67–0.93	
	3	0.88	0.76–1.02	
Marital status[h]				0.01
Single versus married/cohabitating	1	0.43	0.18–0.99	
	2	0.33	0.16–0.67	
	3	0.59	0.31–1.11	
QIDS-C[i]				<0.001
	1	0.56	0.49–0.65	
	2	0.64	0.56–0.72	
	3	0.74	0.66–0.83	
Treatment allocation				0.56
Aug-ARI[j] versus Switch-BUP[k]	1	1.62	0.64–4.12	
	2	1.18	0.54–2.58	
	3	0.98	0.48–1.99	
Aug-BUP[l] versus Switch-BUP	1	0.93	0.38–2.30	
	2	0.69	0.32–1.46	
	3	0.63	0.32–1.26	

[a] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated Core Atypical Cluster. Possible scores are based upon a sum of QIDS-C scores for the 4 following items: psychomotor agitation, psychomotor slowing, suicidal ideation, and hypersomnia. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. The total range is from 0 to 12, with higher scores indicating greater severity of symptoms.
[b] OR, Odds ratio. Odds of inclusion into a specific trajectory in comparison to the odds of inclusion into trajectory 4 (least responsive trajectory).
[c] Age, in years, at baseline.
[d] BAI, Beck Anxiety Inventory, at baseline. Possible scores range from 0 to 3 (average rating of each of the 21 items), with higher scores indicating greater anxiety.
[e] CIRS Severity Index, Cumulative Illness Rating Scale Comorbidity Severity Index, at baseline. Possible scores range from 0 to 4, with higher scores indicating greater severity of co-occurring medical conditions.
[f] Early improvement. The presence of a ≥20% drop from the baseline QIDS-C score by the end of week 2.
[g] Lifetime suicidal ideation, C-SSRS, Columbia Suicide Severity Rating Scale-Lifetime Suicidal Ideation, at baseline. Possible scores range from 0 to 5, with higher scores indicating greater suicidal ideation or intent.
[h] Marital Status, at baseline. Identification of status as single versus married/cohabiting.
[i] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated, at baseline. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of 3 indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27.
[j] Aug-ARI, allocation to augmentation with aripiprazole.
[k] Switch-BUP, allocation to switching to bupropion.
[l] Aug-BUP, allocation to augmentation with bupropion.

and should be cautiously interpreted because of the predominance of Caucasians in this analytical sample. Finally, the presence of lifetime suicidal ideation expectedly appears to influence the magnitude of the QIDS-C atypical symptom cluster, which includes rating of acute suicidal ideation.

Strengths and Limitations

The present analysis of the VAST-D data has several strengths. First, because of the large patient population, this study was ideally suited to perform trajectory analysis and obtain precise estimates of the effect of various factors on trajectory membership. Second, the present analysis

TABLE 7. Exploratory comparisons of demographic and Clinical characteristics among trajectory groups for the total QIDS-C.[a], individual QIDS-C clusters, and the PHQ-9.[b]

	QIDS-C total ($F_{5,999}$)		QIDS-core[c] emotional cluster ($F_{3,999}$)		QIDS-sleep[d] cluster ($F_{3,999}$)		QIDS-atypical[e] cluster ($F_{3,999}$)		PHQ-9 ($F_{5,999}$)	
Categorical factors	Chi-square	*p* value	Chi-square	*p* value	Chi-square	*p* value	Chi-square	*p* value	Chi-square	*p* value
Education[f]	14.8	0.46	7.72	0.56	22.8	0.09	9.87	0.36	13.1	0.59
Employment status[g]	36.5	<0.001	30.5	<0.001	31.2	<0.001	18.5	0.01	23.9	0.01
Gender[h]	10.9	0.05	4.77	0.19	12.6	0.03	21.1	<0.001	24.0	<0.001
Grief endorsement[i]	18.6	<0.001	22.8	<0.001	27.3	<0.001	1.44	0.70	26.9	<0.001
Marital status[j]	1.78	0.87	1.5	0.59	11.4	0.04	7.8	0.05	6.7	0.24
Race[k]	23.6	0.01	8.67	0.19	49.3	<0.001	9.46	0.15	14.8	0.14
Substance or alcohol abuse[l]	4.7	0.45	0.67	0.88	2.1	0.84	6.44	0.09	4.52	0.48
Treatment allocation[m]	32.5	<0.001	16.4	0.01	8.3	0.60	9.31	0.16	8.97	0.54
Continuous factors	*F* statistic ($F_{5,999}$)	*p* value	*F* statistic ($F_{3,999}$)	*p* value	*F* statistic ($F_{3,999}$)	*p* value	*F* statistic ($F_{3,999}$)	*p* value	*F* statistic ($F_{5,999}$)	*p* value
ACES[n]	1.61	0.15	1.20	0.31	2.0	0.08	0.98	0.40	1.03	0.40
Age[o]	1.18	0.32	1.64	0.18	2.4	0.04	2.03	0.11	1.23	0.29
BAI[p]	33.5	<0.001	40.3	<0.001	26.0	<0.001	28.5	<0.001	60.8	<0.001
CIRS severity index[q]	0.18	0.97	1.09	0.35	0.8	0.56	0.99	0.40	1.12	0.35
DSM-5 mixed features[r]	2.13	0.06	0.36	0.79	4.2	<0.001	4.08	0.01	3.15	0.01
Duration of index episode[s]	5.66	<0.001	5.35	<0.01	3.7	<0.001	4.20	0.01	5.61	<0.001
Lifetime suicidal ideation[t]	2.94	0.012	1.14	0.33	0.7	0.59	8.64	<0.001	1.25	0.28
QIDS-C[u]	130	<0.001	141	<0.001	33.0	<0.001	70.9	<0.001	235	<0.001
Q-LES-Q-SF[v]	57.7	<0.001	89.5	<0.001	23.3	<0.001	20.4	<0.001	99.0	<0.001

[a] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of 3 indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27. Comparisons among six trajectory groups (Supporting Information S1: Supplement F, Table 1a).

[b] PHQ-9, Patient Health Questionnaire-9. Each of nine questions is scored from 0 to 3, with 3 indicating greater severity. Possible scores on the PHQ-9 range from 0 to 27, with higher scores indicating greater degree of depression. Comparisons among six trajectory groups (Supporting Information S1: Supplement F, Table 1e).

[c] QIDS-C Core Emotional Cluster, Quick Inventory of Depressive Symptomatology–Clinician Rated Core Emotional Cluster. Possible scores are based upon a sum of QIDS-C scores for the 5 following items: energy/fatigability, concentration/decision making, loss of interest, mood, and feelings of worthlessness. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. Scores range from 0 to 15, with higher scores indicating greater severity of symptoms. Comparisons among four trajectory groups (Supporting Information S1: Supplement F, Table 1b).

[d] QIDS-C Sleep Cluster, Quick Inventory of Depressive Symptomatology–Clinician Rated Sleep Cluster. Possible scores are based upon a sum of QIDS-C scores for the following three items: mid-nocturnal insomnia, sleep-onset insomnia, and early morning insomnia. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. Scores range from 0 to 9, with higher scores indicating greater severity of symptoms. Comparisons among six trajectory groups (Supporting Information S1: Supplement F, Table 1c).

[e] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated Core Atypical Cluster. Possible scores are based upon a sum of QIDS-C scores for the 4 following items: psychomotor agitation, psychomotor slowing, suicidal ideation, and hypersomnia. Each item is rated on a scale from 0 to 3, with 3 representing greater severity. The total range is from 0 to 12, with higher scores indicating greater severity of symptoms. Comparisons among four trajectory groups (Supporting Information S1: Supplement F, Table 1d).

[f] Education, at baseline. The level of educational attainment by the following categories: high school or less, some college, associate degree, bachelor degree, or higher (Supporting Information S1: Supplement F, Tables 1a–1e).

[g] Employment status, at baseline. The employment status by the following categories: unemployed (includes disability or assistance), retired (and not working), or employed (Supporting Information S1: Supplement F, Tables 1a–1e).

[h] Gender. Male or female gender (Supporting Information S1: Supplement F, Tables 1a–1e).

[i] Grief endorsement, at baseline. Endorsement of ≤3 versus >3 items on the Complicated Grief Questionnaire, with endorsement of more items indicating greater complicated grief (Supporting Information S1: Supplement F, Tables 1a–1e).

[j] Marital status, at baseline. Identification of status as single versus married/cohabiting. (Supporting Information S1: Supplement F, Tables 1a–1e).

[k] Race. The declared race of the participant in the following categories: white, African-American/black, or other (Supporting Information S1: Supplement F, Tables 1a–1e).

[l] Substance or alcohol abuse, at baseline. The presence of a substance or alcohol abuse diagnosis by the M.I.N.I. (Supporting Information S1: Supplement F, Tables 1a–1e).

[m] Treatment allocation. Allocation to one of three treatment groups: Aug-ARI, Aug-BUP, or Switch-BUP (Supporting Information S1: Supplement F, Tables 1a–1e).

[n] ACES, Adverse Childhood Experiences Survey, at baseline. Possible scores range from 0 to 10, with higher scores indicating greater childhood adversity and greater risk of psychological or health problems (Supporting Information S1: Supplement G, Tables 1a–1e).

[o] Age, in years, at baseline (Supporting Information S1: Supplement F, Tables 1a–1e).

[p] BAI, Beck Anxiety Inventory, at baseline. Possible scores range from 0 to 3 (average rating of each of the 21 items), with higher scores indicating greater anxiety (Supporting Information S1: Supplement F, Tables 1a–1e).

extracted trajectories based on outcomes over time, without regard to treatment allocation, minimizing any associated bias. Third, we explored the role of novel factors such as duration of index episode, childhood adversity, and quality of life in predicting trajectory group membership. Finally, we used commonly used tools of psychometric measurement for depression scores, which improves the generalizability of the findings and facilitates comparisons with other studies.

This study does have its limitations. First, even though the VAST-D trial was conducted in a diverse sample with regard to many demographics and historical features (38), the gender of the patient population was predominantly male (approximately 85%), and the race of participants was predominantly Caucasian (approximately 70%). Hence, the findings of effects of gender and race on trajectory group membership in the exploratory analysis should be interpreted with caution. A second limitation is that trajectories are only approximations of possible patterns in the population. Some argue that GBTM may "overextract" trajectories, resulting in too many complicated patterns and leading to uncertain clinical conclusions (4, 39–41). Third, because individuals are assigned trajectory groups based on their posterior probability of membership, some uncertainty in trajectory membership may exist. In this analysis, we attempted to address this uncertainty in the multinomial regression by using posterior probability of trajectory group membership as a weight in multinomial logistic regression analysis. Finally, it is also possible that we may have undervalued the role of factors with potential to influence treatment response (e.g. familial history of depression or number of life stressors) that were not studied.

Importance of Findings

In this study, we used the VAST-D trial database to explore differential trajectories of improvement in depression scores using a relatively new statistical procedure, GBTM (3–5, 40). Our characterization of the response trajectories helps to establish reasonable short-term clinical expectations of an antidepressant trial. Although we identified several factors that were associated with specific patterns of response, across multiple statistical probes, we consistently recognized three factors: baseline depression, baseline anxiety, and early improvement. Thus, GTBM analysis based on the large population of the VAST-D data set findings (N = 1522), substantiated a previous GBTM study with a much smaller number of advanced age participants (N = 453) (21) identifying that baseline depression and anxiety severity are very important factors in determining outcome. In addition, our weighted multinomial logistic regression analysis following GTBM corroborates another VAST-D analysis highlighting the role of baseline depression and anxiety in determining response (23). Our GTBM analyses also confirmed the role of early improvement (by the end of Week 2) as a predictor of response (33). Finally, the PHQ-9, as a self-report measure, reproduced the key findings of the QIDS-C, reinforcing the benefit of using it to follow antidepressant response in clinical settings.

AUTHOR AND ARTICLE INFORMATION

Department of Psychiatry, Baylor Scott & White Health, Temple, Texas (Hicks); Texas A&M College of Medicine, Temple, Texas (Hicks); Biostatistics Service, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York (Sevilimedu); Yale University School of Public Health, New Haven, Connecticut (Sevilimedu); Cooperative Studies Program Coordinating Center, VA Connecticut Healthcare System, West Haven, Connecticut (Sevilimedu, Johnson); VA San Diego Healthcare System, San Diego, California (Tal, Zisook); Department of Psychiatry, VISN10 Geriatric Research, Education and Clinical Center, VA Northeast Ohio Healthcare System, Cleveland, Ohio (Chen); Case Western Reserve University, Cleveland, Ohio (Chen); Tuscaloosa VA Medical Center, Tuscaloosa, Alabama (Davis); University of Alabama School of Medicine, Birmingham, Alabama (Davis); Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, Albuquerque, New Mexico (Vertrees); University of California, San Diego, California (Zisook); Veterans Affairs (VA) New England Mental Illness Research, Education and Clinical Center, VA Connecticut Healthcare System, West Haven, Connecticut (Mohamed); Yale University School of Medicine, New Haven, Connecticut (Mohamed).

Send correspondence to Dr. Hicks (Paul.Hicks@BSWHealth.org).

Paul B. Hicks and Varadan Sevilimedu contributed equally to this paper.

Previous Presentation: Components of the data presented in this article were discussed at the annual meeting of the American Psychiatric Association, May 5–9, 2018, New York City.

This study was supported and conducted by the Cooperative Studies Program (CSP 576), Department of Veterans Affairs, Office of Research and Development. The CSP was involved in the design and conduct of the study; the collection, management, analysis, and interpretation of

[q] CIRS Severity Index, Cumulative Illness Rating Scale Comorbidity Severity Index, at baseline. Possible scores range from 0 to 4, with higher scores indicating greater severity of co-occurring medical conditions (Supporting Information S1: Supplement F, Tables 1a–1e).

[r] DSM-5 mixed features, presence of mixed features by a self-rated 9-item mixed features scale based on the DSM-5, at baseline. Possible scores range from 0 to 18, with higher scores indicating more hypomanic or manic symptoms (Supporting Information S1: Supplement F, Tables 1a–1e).

[s] Duration of index episode, duration in months of the depression episode that is currently being treated, at baseline (Supporting Information S1: Supplement F, Tables 1a–1e).

[t] Lifetime Suicidal Ideation, Columbia Suicide Severity Rating Scale-Lifetime Suicidal Ideation, at baseline. Possible scores range from 0 to 5, with higher scores indicating greater suicidal ideation or intent (Supporting Information S1: Supplement F, Tables 1a–1e).

[u] QIDS-C, Quick Inventory of Depressive Symptomatology–Clinician Rated, at baseline. The QIDS-C score is calculated from a total of 16 clinician-administered questions, which map onto 9 psychiatric domains: sleep, mood, appetite, concentration, guilt, acute suicidal ideation, interest, fatigue, and psychomotor function. Each domain is scored on a scale of 0–3, with a score of three indicating greater severity. These scores are then added up to obtain a total QIDS-C score ranging from 0 to 27 (Supporting Information S1: Supplement F, Tables 1a–1e).

[v] Q-LES-Q-SF, Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form, at baseline. Possible scores range from 0% to 100% of the maximum scale score of 70, with higher scores indicating greater life satisfaction and enjoyment (Supporting Information S1: Supplement F, Tables 1a–1e).

the data; and the preparation, review, and approval of the manuscript. The CSP had no role in the decision to submit the manuscript for publication. Bristol-Myers Squibb provided aripiprazole (Abilify) for use in this study.

The authors thank the local site investigators, independent evaluators, nurse coordinators, and patient participants at the 35 VAST-D enrollment sites (listed in Supporting Information S1: Supplement A), as well as the CSP Coordinating Center at the VA Connecticut Healthcare System, West Haven, for providing statistical analyses.

Disclaimer: The opinions expressed in this article are those of the authors and do not necessarily represent the views of the U.S. Department of Veterans Affairs or the U.S. government.

Mr. Johnson owns stock in Bristol-Myers Squibb, where his spouse is an employee. Dr. Davis has received research funding from Tonix and Merck as well as personal consulting fees from Bracket, Janssen, Otsuka, Lundbeck, and Tonix. Dr. Zisook receives funding from Defender Pharmaceuticals and COMPASS Pathways. Dr. Chen has received royalties from UpToDate. The other authors report no financial relationships with commercial interests.

Clinicaltrials.gov identifier: NCT01421342.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. Psychiatric Research and Clinical Practice published by Wiley Periodicals LLC on behalf of American Psychiatric Association.

Received April 17, 2023; revised June 12, 2023; accepted June 29, 2023.

REFERENCES

1. Murray CJ, Atkinson C, Bhalla K, Birbeck G, Burstein R, Chou D, et al. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. JAMA. 2013;310(6):591–608. https://doi.org/10.1001/jama.2013.13805
2. Smagula SF, Butters MA, Anderson SJ, Lenze EJ, Dew MA, Mulsant BH, et al. Antidepressant response trajectories and associated clinical prognostic factors among older adults. JAMA Psychiatr. 2015;72(10):1021–1028. https://doi.org/10.1001/jamapsychiatry.2015.1324
3. Uher R, Mors O, Rietschel M, Rajewska-Rager A, Petrovic A, Zobel A, et al. Early and delayed onset of response to antidepressants in individual trajectories of change during treatment of major depression: a secondary analysis of data from the Genome-Based Therapeutic Drugs for Depression (GENDEP) study. J Clin Psychiatry. 2011;72(11):1478–1484. https://doi.org/10.4088/jcp.10m06419
4. Mohamed S, Johnson GR, Chen P, Hicks PB, Davis LL, Yoon J, et al. Effect of antidepressant switching vs augmentation on remission among patients with major depressive disorder unresponsive to antidepressant treatment: the VAST-D randomized clinical trial. JAMA. 2017;318(2):132–145. https://doi.org/10.1001/jama.2017.8036
5. Nagin DS, Odgers CL. Group-based trajectory modeling (nearly) two decades later. J Quant Criminol. 2010;26(4):445–453. https://doi.org/10.1007/s10940-010-9113-7
6. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol. 2010;6(1):109–138. https://doi.org/10.1146/annurev.clinpsy.121208.131413
7. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. Stat Methods Med Res. 2016;1(7):1–9. https://doi.org/10.1177/0962280216673085
8. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatr. 2003;54:573–583.
9. Mohamed S, Johnson GR, Vertrees JE, Guarino PD, Weingart K, Young IT, et al. The VA augmentation and switching treatments for improving depression outcomes (VAST-D) study: rationale and design considerations. Psychiatr Res. 2015;229(3):760–770. https://doi.org/10.1016/j.psychres.2015.08.005
10. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-IV-TR ed. Washington: American Psychiatric Association; 2000.
11. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x
12. Wisniewski SR, Rush AJ, Balasubramani GK, Trivedi MH, Nierenberg AA; STARD Investigators. Self-rated global measure of the frequency, intensity, and burden of side effects. J Psychiatr Pract. 2006;12(2):71–79. https://doi.org/10.1097/00131746-200603000-00002
13. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998;59(Suppl 20):22–33; quiz 34–57.
14. Kessler RC, Magee WJ. Childhood adversities and adult depression: basic patterns of association in a US national survey. Psychol Med. 1993;23(3):679–690. https://doi.org/10.1017/s0033291700025460
15. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol. 1988;56(6):893–897. https://doi.org/10.1037/0022-006x.56.6.893
16. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatr. 2011;168(12):1266–1277. https://doi.org/10.1176/appi.ajp.2011.10111704
17. Shear KM, Jackson CT, Essock SM, Donahue SA, Felton CJ. Screening for complicated grief among Project Liberty service recipients 18 months after September 11, 2001. Psychiatr Serv. 2006;57(9):1291–1297. https://doi.org/10.1176/ps.2006.57.9.1291
18. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-V ed. Arlington: American Psychiatric Association; 2013.
19. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. J Am Geriatr Soc. 1968;16(5):622–626. https://doi.org/10.1111/j.1532-5415.1968.tb02103.x
20. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. Psychopharmacol Bull. 1993;29:321–326.
21. Jones LV, Thissen D. A history and overview of psychometrics. In: Rao CR, Sinharay S, editors. Handbook of statistics. The Netherlands: Elsevier; 2006. p. 1–27.
22. SAS/IML 14.1 user's guide. Cary: SAS Institute Inc.; 2015.
23. Zisook S, Johnson GR, Tal I, Hicks P, Chen P, Davis L, et al. General predictors and moderators of depression remission: a VAST-D report. Am J Psychiatr. 2019;176(5):348–357. https://doi.org/10.1176/appi.ajp.2018.18091079
24. Blom MB, Spinhoven P, Hoffman T, Jonker K, Hoencamp E, Haffmans PMJ, et al. Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. J Affect Disord. 2007;104(1-3):119–126. https://doi.org/10.1016/j.jad.2007.03.010

25. Riedel M, Möller HJ, Obermeier M, Adli M, Bauer M, Kronmüller K, et al. Clinical predictors of response and remission in inpatients with depressive syndromes. J Affect Disord. 2011;133 (1-2):137–149. https://doi.org/10.1016/j.jad.2011.04.007
26. De Carlo V, Calati R, Serretti A. Socio-demographic and clinical predictors of non-response/non-remission in treatment resistant depressed patients: a systematic review. Psychiatr Res. 2016;240: 421–430. https://doi.org/10.1016/j.psychres.2016.04.034
27. Friedman ES, Davis LL, Zisook S, Wisniewski SR, Trivedi MH, Fava M, et al. Baseline depression severity as a predictor of single and combination antidepressant treatment outcome: results from the CO-MED trial. Eur Neuropsychopharmacol. 2012;22(3):183–199. https://doi.org/10.1016/j.euroneuro.2011.07.010
28. Papakostas GI, Fan H, Tedeschini E. Severe and anxious depression: combining definitions of clinical sub-types to identify patients differentially responsive to selective serotonin reuptake inhibitors. Eur Neuropsychopharmacol. 2012;22(5):347–355. https://doi.org/10.1016/j.euroneuro.2011.09.009
29. Dold M, Bartova L, Souery D, Mendlewicz J, Serretti A, Porcelli S , et al. Clinical characteristics and treatment outcomes of patients with major depressive disorder and comorbid anxiety disorders — results from a European multicenter study. J Psychiatr Res. 2017;91:1–13. https://doi.org/10.1016/j.jpsychires.2017.02.020
30. Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. Am J Psychiatr. 2008;165(3):342–351. https://doi.org/10.1176/appi.ajp.2007.06111868
31. Trivedi MH, Rush AJ, Wisniewski SR, Warden D, McKinney W, Downing M, et al. Factors associated with health-related quality of life among outpatients with major depressive disorder: a STAR*D report. J Clin Psychiatry. 2006;67(02):185–195. https://doi.org/10.4088/jcp.v67n0203
32. Gorwood P, Bayle F, Vaiva G, Courtet P, Corruble E, Llorca PM. Is it worth assessing progress as early as week 2 to adapt antidepressive treatment strategy? Results from a study on agomelatine and a global meta-analysis. Eur Psychiatr. 2013;28(6):362–371. https://doi.org/10.1016/j.eurpsy.2012.11.004
33. Hicks PB, Sevilimedu V, Johnson GR, Tal I, Chen P, Davis LL, et al. Predictability of nonremitting depression after first 2 weeks of antidepressant treatment: a VAST-D trial report. Psychiatr Res Clin Pract. 2019;1:58–67. https://doi.org/10.1176/appi.prcp.20190003
34. Olgiati P, Serretti A, Souery D, Dold M, Kasper S, Montgomery S, et al. Early improvement and response to antidepressant medications in adults with major depressive disorder. Meta-analysis and study of a sample with treatment-resistant depression. J Affect Disord. 2018;227:777–786. https://doi.org/10.1016/j.jad.2017.11.004
35. 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. https://doi.org/10.4088/jcp.07m03780
36. Wagner S, Engel A, Engelmann J, Herzog D, Dreimüller N, Müller MB, et al. Early improvement as a resilience signal predicting later remission to antidepressant treatment in patients with major depressive disorder: systematic review and meta-analysis. J Psychiatr Res. 2017;94:96–106. https://doi.org/10.1016/j.jpsychires.2017.07.003
37. Chekroud AM, Gueorguieva R, Krumholz HM, Trivedi MH, Krystal JH, McCarthy G. Reevaluating the efficacy and predictability of antidepressant treatments: a symptom clustering approach. JAMA Psychiatr. 2017;74(4):370–378. https://doi.org/10.1001/jamapsychiatry.2017.0025
38. Zisook S, Tal I, Weingart K, Hicks P, Davis LL, Chen P, et al. Characteristics of U.S. veteran patients with major depressive disorder who require "next-step" treatments: a VAST-D report. J Affect Disord. 2016;206:232–240. https://doi.org/10.1016/j.jad.2016.07.023
39. Pines HA, Gorbach PM, Weiss RE, Shoptaw S, Landovitz RJ, Javanbakht M, et al. Sexual risk trajectories among MSM in the United States: implications for pre-exposure prophylaxis delivery. J Acquir Immune Defic Syndr. 2014;65(5):579–586. https://doi.org/10.1097/qai.0000000000000101
40. Bauer DJ, Curran PJ. Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. Psychol Methods. 2003;8(3):338–363. https://doi.org/10.1037/1082-989x.8.3.338
41. Nagin DS, Tremblay RE. Developmental trajectory groups: fact or a useful statistical fiction? Criminology. 2005;43(4):873–904. https://doi.org/10.1111/j.1745-9125.2005.00026.x