Title: Longitudinal Trajectories of Symptom Change During Antidepressant Treatment Among Managed Care Patients with Co-Occurring Depression and Anxiety

Authors and Affiliations:

1. Judith Cukor, PhD  Weill Cornell Medicine (New York, NY, USA)
2. Zhenxing Xu, PhD  Weill Cornell Medicine (New York, NY, USA)
3. Veer Vekaria, BS  Weill Cornell Medicine (New York, NY, USA)
4. Fei Wang, PhD  Weill Cornell Medicine (New York, NY, USA)
5. Mark Olfson, MD  Columbia University (New York, NY, USA)
6. Samprit Banerjee, PhD  Weill Cornell Medicine (New York, NY, USA)
7. Gregory Simon, MD  Kaiser Permanente Washington Health Research Institute (Seattle, WA, USA)
8. George Alexopoulos, MD  Weill Cornell Medicine (New York, NY, USA)
9. Jyotishman Pathak, PhD  Weill Cornell Medicine (New York, NY, USA)

Correspondence and Offprint Requests:

Jyotishman Pathak, PhD
425 East 61st Street, 3rd Floor, New York, NY 10065
+1 (646) 962-9415
jyp2001@med.cornell.edu

Keywords:  Depression, Anxiety, Antidepressant, Electronic Health Records, Machine Learning
Longitudinal Trajectories of Symptom Change During Antidepressant Treatment
Among Managed Care Patients with Co-Occurring Depression and Anxiety

ABSTRACT: Depression and anxiety are highly correlated, yet little is known about the course of each condition when presenting concurrently. This study aimed to identify longitudinal patterns and changes in depression and anxiety symptoms during antidepressant treatment, and evaluate clinical factors associated with each response pattern. Self-report Patient Health Questionnaire-9 (PHQ-9) and General Anxiety Disorder-7 (GAD-7) scores were used to track the courses of depression and anxiety respectively over a three-month window, and group-based trajectory modeling was used to derive subgroups of patients who have similar response patterns. Multinomial regression was used to associate various clinical variables with trajectory subgroup membership. Of the 577 included adults, 373 (64.6%) were women, and the mean age was 39.3 (SD: 12.9) years. Six depression and six anxiety trajectory subgroups were computationally derived; three depression subgroups demonstrated symptom improvement, and three exhibited nonresponse. Similar patterns were observed in the six anxiety subgroups. Factors associated with treatment nonresponse included higher pretreatment depression and anxiety severity and poorer sleep quality, while better overall health and younger age were associated with higher rates of remission. Synchronous and asynchronous paths to improvement were also observed between depression and anxiety. High baseline depression or anxiety severity alone may be an insufficient predictor of treatment nonresponse. These findings have the potential to motivate clinical strategies aimed at treating depression and anxiety simultaneously.

Key Words: Depression, Anxiety, Antidepressant, Electronic Health Records, Machine Learning
INTRODUCTION

Symptoms of depression and anxiety often coexist. Anxious depression accounts for 42-78% of all depression cases and has been associated with higher immune and HPA-axis dysregulations and corticolimbic connectivity and more prominent cortical thinning than non-anxious depression [1]. Overall, anxious depression has poorer clinical outcomes than non-anxious depression [2]. Treatment with antidepressants is common, yet response to antidepressant treatment remains inconsistent, and dosage or medication changes often occur on a trial and error basis [3, 4].

Previous research has begun to explore trajectories of response to depression or anxiety treatment over time [5-10]. However, studies of comorbid symptom response patterns are largely limited to psychotherapy treatments [5, 7, 9], and despite the high correlation and increased functional impairment in patients with comorbid anxiety and depression symptoms, little remains known about the course of each condition when presenting together in the context of antidepressant treatment. Furthermore, many studies do not take into account the intermediate oscillations in patients’ symptoms throughout the course of the treatment period, thereby misrepresenting differences in trajectories that may exist.

One study by Smagula et al. [8] analyzed data on antidepressant response in a population of older adults from a 12 week pharmacological treatment study identified 6 distinct trajectories, including 3 responding and 3 nonresponding. Responding patients had rapid or delayed responses. A second study by Sunderland et al. [9] investigated trajectories of change for distinct depressed and anxious patients receiving internet-based cognitive behavioral therapy identified two groups with higher and lower response rates across sessions. Another study by Saunders et al. [7] investigated changes in symptoms of depression and anxiety over the course of psychotherapy identified 4 distinct trajectories of treatment response for depressive symptoms, and 5 trajectories for anxiety symptoms, including some with a delayed response. A large overlap was seen in terms of depressive and anxiety improvement, especially among those who did not respond to treatment for either diagnosis.

The most commonly identified predictor of poorer treatment response was higher baseline levels
of symptom severity [6-9]. Other predictors included psychological distress [9], older age, poorer self-reported physical health and unemployment [6], longer duration of symptoms, better reported sleep, more guilt, more functional impairment [8], lower socioeconomic status, and poorer overall functioning [7].

Providing feedback to clinicians about their patients’ anticipated symptom change using expected response trajectories can improve outcomes [7], and medication decisions are informed by anticipated responses. However, the usefulness of expected trajectories may be limited if anticipated outcomes are modeled upon an averaged response curve for patients. This may reduce the precision of clinical predictions and diminish the quality of management decisions.

Understanding clinical response trajectories and their contributing factors can be instrumental in beginning to predict clinical response. It may further serve as a significant step toward personalized medicine. Compared to traditional metrics for depression which neither consider variability across multiple time points nor account for co-occurring anxiety, understanding the course of response for depression and anxiety and how they relate to each other can help further elucidate the relationship with each other [11].

Group-based trajectory modeling (GBTM), a latent class analysis method, provides a novel application for evaluating treatment response variability without depending on a prespecified remission threshold, in contrast to some prevailing literature [8, 11]. GBTM is a data-driven method that allows identification of groups of patients that share common characteristics such as demographics, comorbidities, and outcomes including treatment response [12]. Previous studies on depression have demonstrated that GBTM has the potential to capture heterogeneity during the course of illness and recovery [13, 14].

The aims of study were to identify (1) patterns (trajectories) of symptom change and (2) patient and clinical characteristics associated with increased likelihood of following the identified trajectories. To achieve this, we applied GBTM to capture the typical trajectories of symptom change throughout the initial course of antidepressant treatment in patients with co-occurring depression and anxiety. A heatmap was generated to analyze the synchronous and asynchronous
relationships between the trajectories of depression and anxiety to delineate what clinicians are likely to observe in practice. In addition, the associations of baseline depression and anxiety symptom severity, sleep quality, characteristics, medical comorbidity, and overall health status with trajectory group membership were examined to identify potential predictors of symptom improvement.

METHODS

Overview
This cohort study used electronic health records (EHRs) and repeated Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder (GAD-7) measurements in 577 patients who received outpatient mental healthcare services and treatment at an urban academic medical center. We used GBTM to identify patients with similar trajectories of PHQ-9 and GAD-7 change following initiation of antidepressant therapy. These trajectory groups were then characterized using comprehensive patient information from their EHRs including demographics, comorbidities, and a variety of scoring instruments and questionnaires. The synchronicity of PHQ-9 and GAD-7 trajectories for individual patients was then described. Finally, a multinomial regression model was built to investigate the associations of prognostic factors with trajectory group membership. The overall study design is illustrated in Figure 1.

Cohort Description
Individual-level EHR data were obtained from the outpatient behavioral health clinic at a major medical center in New York City. Patients at this clinic were administered multiple mental health and overall physical health related questionnaires as part of routine clinical encounters. Patients (n=577) with multiple self-reported PHQ-9 (n=3,730) and GAD-7 (n=3,704) scores between June 13, 2016 and September 26, 2020 were included. Adult patients (≥18 years old) were included in this study if a) they received antidepressant prescriptions 3 months before or up to one week after establishing care at the clinic, and b) both PHQ-9 and GAD-7 scores were measured at baseline. Prescribed antidepressants included selective serotonin reuptake inhibitors (SSRIs) (63.95%), atypical antidepressants (22.18%), serotonin and norepinephrine reuptake inhibitors (SNRIs) (9.53%), tricyclic antidepressants (TCAs) (3.99%), and monoamine oxidase inhibitors (MAOIs) (0.35%).
To ensure sufficient detail to model trajectories of depression and anxiety symptom change, patients (n=1,069) with fewer than three treatment sessions were excluded. For other patients, missing values regarding PHQ-9 and GAD-7 scores during a three-month period were imputed using the last observation carried forward (LOCF) or first observation carried backward (FOCB) method. Since a PHQ-9 score of $\geq 10$ and GAD-7 score of $\geq 8$ represents a reasonable cut-off for identifying probable cases of depressive disorder [7] and generalized anxiety disorder [15], respectively, patients were also excluded (n=726) if their baseline assessments were PHQ-9 <10 or GAD-7 <8. The inclusion/exclusion cascade for identifying study participants is shown in eFigure 1. IRB exemption was granted by the Weill Cornell Medicine (WCM) Institutional Review Board (IRB).

**Measures**

**Outcomes**

Depression symptom severity was measured with the PHQ-9 using symptom severity ranges of absent (0-4), mild (5-9), moderate (10-14), moderately severe (15-19), or severe (20-27) [13]. Anxiety symptom severity was measured with the GAD-7 using symptom severity ranges of minimal (0-4), mild (5-9), moderate (10-14), or severe (15-21) [14]. Response was defined as a follow-up PHQ-9 (or GAD-7) score that was at least 50% lower than the patient’s baseline score [6]. Remission was defined as achieving a follow-up score of $\leq 5$ [6].

**Baseline characteristics**

Baseline characteristics include demographic (age, gender, and race), comorbidities (post-traumatic stress disorder (PTSD), bipolar disorder, and “suicidal behavior”), and multiple validated scoring instruments and questionnaires (Alcohol Use Disorders Identification Test (AUDIT) [16], Columbia-Suicide Severity Rating Scale (C-SSRS) [17]; Pittsburgh Sleep Quality Index (PSQI) [18]; Patient-Reported Outcomes Measurement Information System (PROMIS) [19], and Sheehan Disability Scale (SDS) [20]. More details about these baseline characteristics are included in the supplementary materials.

**Statistical Analysis**
Analysis was conducted in R (version 4.0.3) and GBTM [21] was performed using the package lcmm [22]. The number of groups was determined by using the Akaike Information Criterion (AIC) [23] and Bayesian Information Criterion (BIC) [24]. Lower AIC and BIC values for one model compared to another indicate better model fit. In addition, following convention with GBTM models, each class needed to contain at least 5% of the sample for it to be considered meaningful and numerically stable [7][8]. The univariate and multivariate tests were performed to obtain associations between baseline prognostic factors and trajectory groups. For the univariate analyses, we applied a student's t-test, Mann-Whitney U test, Chi-square test, and Fisher’s exact test where appropriate. For the multivariate analyses, multinomial logistic regression models were used.

RESULTS

Characteristics of the derived trajectory groups

Six trajectory groups representing depressive symptom change were derived and named according to their relative baseline score and response pattern. Additional details about the process for choosing the optimal number of trajectory groups in GBTM using AIC and BIC are shown in the supplementary file (eTables 3 and 4). These included “high, nonresponse” (Cluster 1), “high, response” (Cluster 2), “moderately severe, nonresponse” (Cluster 3), “moderately severe, response” (Cluster 4), “moderate, remission” (Cluster 5), and “moderate, nonresponse” (Cluster 6). With the baseline PHQ-9 score serving as a reference point, these trajectory groups can be roughly partitioned into “severe,” “moderately severe,” and “moderate” strata, consisting of n=111 (19.24%), n=230 (39.86%), and n=236 (40.90%) patients, respectively. Within each stratum, one response and one nonresponse trajectory were observed. Within the “severe” stratum (baseline PHQ-9 ≥20), patients in the high, response group demonstrated a roughly linearly decreasing curve indicating substantial and sustained improvement in depression symptom severity overall, while high, nonresponse patients maintained a roughly constant PHQ-9 score trajectory, indicating neither improving nor worsening symptomatology. When observing the moderately severe group (baseline PHQ-9=15-19), patients in the moderately severe, response group exhibited an inverse sigmoid-shaped curve that was initially flat, then decreased sharply and gradually tapered off. Patients in the moderately severe, nonresponse group, however, showed little to no improvement. Finally, in the moderate groups (baseline PHQ-9
patients in the moderate, remission group showed an exponentially decreasing curve that reached and sustained a level below typical remission criteria (PHQ-9 ≤ 5) by the latter half of the observation window. In contrast, patients in the moderate, nonresponse group demonstrated only minimal improvement. Patient characteristics for the PHQ-9 trajectory groups are described in Table 1 and showed in Figure 2. The similar six trajectory groups on anxiety symptom change were also found, more details were included in supplementary materials.

The overlap of patients between all permutations of the PHQ-9 and GAD-7 trajectory groups was also explored (Figure 3). The pairing with the largest degree of overlap (n=49) occurred between the PHQ-9 moderate, remission group (Cluster 5) and GAD-7 moderate, remission group (Cluster 6), suggesting similar rates of improvement for both depression and anxiety symptoms. The next largest degree of overlap (n=41) occurred between the PHQ-9 moderate, nonresponse group (Cluster 6) and GAD-7 moderate, nonresponse-2 group (Cluster 4), which suggests that a group of patients experience minimal change in depressive and anxiety symptoms during this time period. Analysis of all the overlap permutations revealed that most overlap sets (72.45% patients) displayed similar trendlines in terms of PHQ-9 and GAD-7 score change, which suggests that patients with comorbid depression and anxiety generally experience synchronous trajectories treatment response or nonresponse. However, patients in some overlap sets (27.55% patients) did demonstrate asynchronous symptom change. For example, the overlap set that contains the PHQ-9 moderately severe, nonresponse group (Cluster 3) and GAD-7 high, response group (Cluster 2) exhibited improvement in anxiety symptoms with antidepressant treatment, but similar significant improvements were not observed in depressive symptoms.

**Associations of baseline patient characteristics with trajectory group membership**

Associations of patients’ baseline characteristics with their derived PHQ-9 and GAD-7 trajectory classes were investigated using a multinomial regression model (see Table 2 (PHQ-9) and eTable 2 (GAD-7)). The following are probabilities in relation to being in PHQ-9 trajectory Cluster 5 (Moderate, remission): the probability of being in PHQ-9 trajectory Cluster 1 (High, nonresponse) and Cluster 2 (High, response) were higher and was associated with poorer overall health and more severity of depression prior to treatment (i.e., lower PROMIS scores and higher baseline PHQ-9 scores). The probability of being in PHQ-9 trajectory Cluster 3 (Moderate...
severe, nonresponse) and Cluster 4 (Moderately severe, response) relative to Cluster 5 (Moderate, remission) were higher and were associated with poorer overall health and more severity of depression prior to treatment (lower baseline PROMIS scores and higher baseline PHQ-9 scores). The probability of being in PHQ-9 trajectory Cluster 6 (Moderate, nonresponse) relative to Cluster 5 (Moderate, remission) was higher and was associated with less severity of depression prior to treatment (lower baseline PHQ-9 scores). The associations of baseline patient characteristics with GAD-7 trajectory group membership were included in the supplementary materials.

**DISCUSSION**

This study investigated short-term trajectories of depression and anxiety symptom severity in a cohort of patients during antidepressant treatment in an outpatient behavioral health clinic. GBTM revealed six distinct depression and anxiety response trajectories. Fewer than half of patients were included in the responder groups for depressive (42.98%) and anxiety (44.89%) related symptoms over weeks 1 to 13 of antidepressant treatment. Baseline variables associated with statistically significant differences between the groups for depression included higher PHQ-9 and C-SSRS scores related to poorer symptom response. Similarly, statistically significant differences between the anxiety groups were associated with higher GAD-7 and PSQI scores related to poorer symptom response. The highest degrees of overlap occurred between PHQ-9 and GAD-7 trajectories that follow synchronous paths (72.55%), suggesting a close relationship between symptoms of depression and anxiety. Several factors measured at baseline were also found to be uniquely associated with some trajectories (e.g., PROMIS, PSQI, SDS, and age).

Higher severity of depression and anxiety symptoms prior to antidepressant initiation was generally associated with lack of change. However, there is a small group of patients who achieved a favorable course towards remission despite initiating antidepressant therapy with anxiety and depression symptoms in the severe ranges. Other groups with moderate symptoms at baseline showed unremarkable responses. This illustrates that baseline severity is only one of a number of factors that is correlated with treatment response. Other variables associated with favorable trajectories include lower sleep health (PSQI score), younger age, lower physical and social wellbeing (PROMIS score), and lower alcohol consumption habits (AUDIT). Preliminary
analysis of several individuals (n=94, 16.29%) who received psychotherapy at the site indicated
similar rates of nonresponse relative to the overall sample, but rates of remission vs. response
trended higher among these individuals.

Among the responder groups, there were differences in the rate of response, with at least one
group demonstrating a slow initial response followed by rapid and sustained improvement in
symptoms beginning in the fourth week. This may have clinical relevance in terms of
determining at what point a treatment regimen should be reconsidered for a non-responding
patient. This finding is juxtaposed with previous studies that suggested clinicians should change
or cease treatment if no remarkable improvement was noted by the third therapy session [25, 26].
Baseline demographic and clinical variables related to the identified trajectories of symptom
change can be used to predict the most likely trajectory group for treatment response, which has
a potential to enhance clinical decision making, optimize resource allocation, and improve
treatment outcomes [7]. Patients who are identified as potential responders to treatment may be
candidates for continuing treatment, while those identified as potential nonresponders may
pursue alternative interventions, such as psychotherapy, at an earlier time point [28].

Another interesting finding emerged from our analysis of patients with overlapping trajectories
for anxiety and depression (Figure 3). Despite the high correlation and increased functional
impairment in patients with comorbid anxiety and depression symptoms, little was known about
the course of each condition when presenting together in the context of antidepressant treatment.
The heatmap analysis showed that most patients had synchronous depression and anxiety
trajectories after initiating treatment. In particular, patients who experienced large degrees of
depressive symptomatic change similarly experienced large degrees of anxiety symptomatic
change. However, there were several asynchronous cases (27.45% patients) where patients who
experienced large degrees of depressive symptomatic change experienced small degrees of
anxiety symptomatic change, and vice versa. One reason may be attributed to the different
dosages of SSRIs sometimes necessary to achieve therapeutic benefits for anxiety as opposed to
depression. However, further exploration of the causes for asynchronous change may provide
some clarity regarding the interplay of depression and anxiety for patients with comorbid
symptoms. A number of baseline variables were associated with symptom change or lack
thereof. Older age was associated with nonresponse, which supports previous work that showed treatment resistance in older populations [27]. Poor physical, mental, and social health at baseline (as measured by PROMIS) were also associated with poorer response during antidepressant treatment. These associations were obtained based on multivariable models suggesting shared variance between these factors. Clinicians may consider these factors in their treatment plan, and more closely follow patients with these characteristics.

The findings of this study may have implications for clinical practice in antidepressant therapy and personalized clinical management of patients suffering from comorbid depression and anxiety. First, our analysis of the associations of pretreatment characteristics with symptom change may aid clinicians in making treatment recommendations. For example, the ability to identify likely nonresponders at the point of a baseline assessment might encourage clinicians and patients to pursue other treatment options earlier in the course of treatment, thereby reducing unnecessary trial-and-error, delays and risk of adverse effects. The findings might also suggest what risk factors clinicians should be more aware of in terms of evaluating treatment response trajectories. The baseline characteristics information can also be used to help anticipate likely symptom change trajectories, which can support clinical decision-making and lead to better treatment outcomes. The trajectories we described may also aid in the identification of patients whose symptoms are not improving, as well as deciding at what point clinicians should intervene or switch antidepressants or consider adjuvant treatment. On a broader scale, understanding these sub-populations can facilitate clinicians in personalizing care and management for these patients. Furthermore, our analysis showed that, although there are many features and symptoms that overlap between depression and anxiety, there are some that are unique to each. Our analysis of the several characteristics associated with the anxiety and depression trajectories derived in this study lays the groundwork for deeper investigation of the interaction and impacts that one disorder has on the other.

The results of this study should be understood within the context of some methodological limitations. First, because the analysis was limited to patients receiving antidepressant medications, it is not possible to attribute the symptom trajectories to antidepressant treatment as distinct from regression to the mean. Second, this study used data from a single outpatient
mental health clinic at an urban academic medical center, which limits our ability to generalize
the model’s utility across multiple different settings and health systems. It will be necessary to
replicate the observed response patterns in other cohorts. Third, although commonly associated
with the use of EHR data across a single health system, patients may have sought concurrent
treatment at sites not affiliated with the health system and therefore not captured within our
EHR. This limits our ability to fully control for or reliably measure the influence of such factors
on treatment response.

CONCLUSION

Six different trajectories of symptom change for depression and anxiety were derived. The
observed trajectories included various paths to symptom response and remission that clinicians
may encounter in practice, and casted two often conflated mental health conditions in their own
lights. While our analysis also examined several factors associated with certain trajectory
responses, deeper understanding of the underlying factors of trajectory-based subgroups may
reveal insights into the mixed pathophysiology of depression and anxiety, and motivate strategies
to personalize therapy for these conditions.

Conflict of Interest: None.

Acknowledgements: This research is funded in part by grants from the US National Institutes of
Health (grant numbers R01 GM105688, R01 MH119177, R01 MH121907, and R01
MH121922).
REFERENCES


TABLE LEGENDS

Table 1. Patient features and characteristics for the PHQ-9 trajectory groups

Table 2. Associations of baseline patient characteristics with PHQ-9 trajectory group membership

FIGURE LEGENDS

Figure 1. The framework of the study. Step 1: PHQ-9 and GAD-7 scores self-reported by patients each week were extracted from two datasets. Step 2: The PHQ-9 and GAD-7 scores trajectory during 14 weeks were constructed. Then, a Group-Based Trajectory Modeling (GBTM) was performed on the top of individual PHQ-9 and GAD-7 score trajectories to identify trajectory groups. Step 3: The characteristics of trajectory groups were analyzed. A multinomial regression model was built to investigate the association of baseline variables with the trajectory groups.

Figure 2. The trajectory groups derived from PHQ-9 and GAD-7 scores

Figure 3. The overlap of patients between the PHQ-9 and GAD-7 trajectory groups
**Table 1.** Patient features and characteristics for the PHQ-9 trajectory groups

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All patients</th>
<th>Cluster 1 (High, nonresponse)</th>
<th>Cluster 2 (High, response)</th>
<th>Cluster 3 (Moderately severe, nonresponse)</th>
<th>Cluster 4 (Moderately severe, response)</th>
<th>Cluster 5 (Moderate, remission)</th>
<th>Cluster 6 (Moderate, nonresponse)</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>39.34 (12.87)</td>
<td>44.31 (13.95)</td>
<td>41.71 (14.39)</td>
<td>38.09 (12.22)</td>
<td>39.57 (12.77)</td>
<td>36.76 (10.22)</td>
<td>39.32 (13.57)</td>
<td>0.019</td>
</tr>
<tr>
<td>PHQ-9, mean (SD)</td>
<td>16.19 (4.26)</td>
<td>20.66 (3.05)</td>
<td>22.53 (2.04)</td>
<td>15.8 (2.74)</td>
<td>18.41 (2.92)</td>
<td>13.63 (3.33)</td>
<td>12.57 (2.17)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GAD-7, mean (SD)</td>
<td>15.35 (3.81)</td>
<td>17.03 (3.47)</td>
<td>17.2 (3.45)</td>
<td>15.61 (3.36)</td>
<td>15.7 (4.0)</td>
<td>14.42 (3.78)</td>
<td>14.08 (3.76)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AUDIT, mean (SD)</td>
<td>3.69 (1.61)</td>
<td>3.73 (1.88)</td>
<td>3.35 (1.53)</td>
<td>3.66 (1.43)</td>
<td>3.61 (1.62)</td>
<td>3.71 (1.53)</td>
<td>3.88 (1.69)</td>
<td>0.180</td>
</tr>
<tr>
<td>C-SSRS, mean (SD)</td>
<td>1.47 (1.2)</td>
<td>1.9 (1.19)</td>
<td>2.45 (1.2)</td>
<td>1.5 (1.26)</td>
<td>1.91 (1.1)</td>
<td>1.08 (1.08)</td>
<td>0.88 (0.78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PROMIS, mean (SD)</td>
<td>13.57 (1.89)</td>
<td>12.6 (1.69)</td>
<td>12.53 (1.74)</td>
<td>13.32 (1.67)</td>
<td>13.26 (1.96)</td>
<td>14.25 (1.86)</td>
<td>14.34 (1.69)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PSQI, mean (SD)</td>
<td>10.9 (2.7)</td>
<td>12.34 (1.91)</td>
<td>12.22 (2.86)</td>
<td>10.88 (2.73)</td>
<td>11.09 (2.61)</td>
<td>10.25 (2.52)</td>
<td>10.14 (2.66)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SDS, mean (SD)</td>
<td>19.56 (4.72)</td>
<td>21.47 (4.52)</td>
<td>22.51 (5.4)</td>
<td>19.98 (4.23)</td>
<td>20.03 (3.93)</td>
<td>18.37 (3.91)</td>
<td>17.75 (5.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gender – Female (%)</td>
<td>373 (64.64)</td>
<td>42 (67.74)</td>
<td>33 (67.35)</td>
<td>85 (62.04)</td>
<td>63 (67.74)</td>
<td>70 (66.04)</td>
<td>80 (61.54)</td>
<td>0.869</td>
</tr>
<tr>
<td>Race – White (%)</td>
<td>313 (54.25)</td>
<td>31 (50.00)</td>
<td>28 (57.14)</td>
<td>74 (54.01)</td>
<td>52 (55.91)</td>
<td>55 (51.89)</td>
<td>73 (56.15)</td>
<td>0.949</td>
</tr>
</tbody>
</table>
### Table 2. Associations of baseline patient characteristics with PHQ-9 trajectory group membership

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PHQ-9 Class 1 (High, nonresponse): RR (95%CI) &amp; p-value</th>
<th>PHQ-9 Class 2 (High, response): RR (95%CI) &amp; p-value</th>
<th>PHQ-9 Class 3 (Moderately severe, nonresponse): RR (95%CI) &amp; p-value</th>
<th>PHQ-9 Class 4 (Moderate, response): RR (95%CI) &amp; p-value</th>
<th>PHQ-9 Class 5 (Moderate, remission): RR (95%CI) &amp; p-value</th>
<th>PHQ-9 Class 6 (Moderate, nonresponse): RR (95%CI) &amp; p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.0 (0.97-1.03), p = .9</td>
<td>0.98 (0.95-1.02), p = .319</td>
<td>0.99 (0.97-1.02), p = .531</td>
<td>1.0 (0.97-1.02), p = .819</td>
<td>Reference</td>
<td>1.02 (1.0-1.05), p = .041</td>
</tr>
<tr>
<td>Gender – Female</td>
<td>0.61 (0.27-1.38), p = .234</td>
<td>0.58 (0.23-1.45), p = .243</td>
<td>0.7 (0.39-1.25), p = .23</td>
<td>0.73 (0.37-1.44), p = .361</td>
<td>Reference</td>
<td>0.97 (0.55-1.7), p = .916</td>
</tr>
<tr>
<td>Race – Non-White</td>
<td>1.0 (0.46-2.17), p = .991</td>
<td>0.59 (0.24-1.43), p = .242</td>
<td>0.84 (0.48-1.47), p = .542</td>
<td>0.8 (0.42-1.52), p = .491</td>
<td>Reference</td>
<td>0.94 (0.55-1.61), p = .832</td>
</tr>
<tr>
<td>AUDIT</td>
<td>1.04 (0.82-1.32), p = .721</td>
<td>0.87 (0.65-1.18), p = .378</td>
<td>0.98 (0.81-1.18), p = .794</td>
<td>0.96 (0.78-1.19), p = .723</td>
<td>Reference</td>
<td>1.06 (0.89-1.26), p = .519</td>
</tr>
<tr>
<td>C-SSRS</td>
<td>0.81 (0.54-1.22), p = .311</td>
<td>1.08 (0.71-1.63), p = .716</td>
<td>0.95 (0.69-1.3), p = .736</td>
<td>1.03 (0.73-1.45), p = .868</td>
<td>Reference</td>
<td>0.87 (0.6-1.25), p = .456</td>
</tr>
<tr>
<td>LEC-5</td>
<td>1.08 (0.97-1.2), p = .181</td>
<td>1.07 (0.93-1.22), p = .355</td>
<td>1.08 (1.0-1.16), p = .05</td>
<td>1.05 (0.96-1.16), p = .27</td>
<td>Reference</td>
<td>1.06 (0.98-1.14), p = .132</td>
</tr>
<tr>
<td>PCL-C</td>
<td>1.02 (0.91-1.14), p = .768</td>
<td>1.02 (0.9-1.15), p = .783</td>
<td>1.07 (0.98-1.16), p = .116</td>
<td>1.0 (0.91-1.1), p = .967</td>
<td>Reference</td>
<td>0.95 (0.88-1.04), p = .258</td>
</tr>
<tr>
<td>PROMIS</td>
<td>0.47 (0.4-0.56), p &lt; .001</td>
<td>0.43 (0.36-0.53), p &lt; .001</td>
<td>0.76 (0.68-0.84), p &lt; .001</td>
<td>0.66 (0.58-0.75), p &lt; .001</td>
<td>Reference</td>
<td>1.07 (0.96-1.18), p = .211</td>
</tr>
<tr>
<td>PSQI</td>
<td>0.95 (0.79-1.12), p = .525</td>
<td>0.86 (0.71-1.05), p = .147</td>
<td>0.94 (0.84-1.06), p = .316</td>
<td>0.91 (0.79-1.05), p = .204</td>
<td>Reference</td>
<td>1.0 (0.9-1.12), p = .983</td>
</tr>
<tr>
<td>SDS</td>
<td>0.88 (0.8-0.97), p = .01</td>
<td>0.88 (0.79-0.98), p = .022</td>
<td>1.01 (0.95-1.09), p = .694</td>
<td>0.94 (0.87-1.02), p = .121</td>
<td>Reference</td>
<td>1.0 (0.94-1.07), p = .904</td>
</tr>
<tr>
<td>PHQ-9</td>
<td>2.08 (1.77-2.44), p &lt; .001</td>
<td>2.44 (2.04-2.93), p &lt; .001</td>
<td>1.26 (1.13-1.4), p &lt; .001</td>
<td>1.67 (1.47-1.91), p &lt; .001</td>
<td>Reference</td>
<td>0.88 (0.78-0.98), p = .021</td>
</tr>
<tr>
<td>GAD-7</td>
<td>0.97 (0.86-1.1), p = .635</td>
<td>0.95 (0.83-1.09), p = .462</td>
<td>1.01 (0.93-1.1), p = .797</td>
<td>0.98 (0.89-1.08), p = .679</td>
<td>Reference</td>
<td>1.03 (0.95-1.12), p = .472</td>
</tr>
</tbody>
</table>
Analysis of Identified Trajectory Groups

Data Preparation

- Extraction of PHQ-9 and GAD-7 scores each week
- Derivation of PHQ-9 and GAD-7 score trajectories during 14 weeks of antidepressant treatment

PHQ-9 and GAD-7 scores measure the severity of depression and anxiety, respectively.

Derivation of Trajectory Groups

Antidepressant prescribed

- Group-based trajectory modeling
- Latent class analysis

Clusters of statistically similar trajectory patterns

Analysis of Identified Trajectory Groups

- The overlap between the identified PHQ-9 and GAD-7 trajectory groups
- The characteristics of trajectory groups in terms of demographics, type of response, and intensity

Demographics

- Associations of baseline variables with group membership using a multinomial regression model

Questionnaires
Figure 1. The framework of the study. Step 1: PHQ-9 and GAD-7 scores self-reported by patients each week were extracted from two datasets. Step 2: The PHQ-9 and GAD-7 scores trajectory during 14 weeks were constructed. Then, a Group-Based Trajectory Modeling (GBTM) was performed on the top of individual PHQ-9 and GAD-7 score trajectories to identify trajectory groups. Step 3: The characteristics of trajectory groups were analyzed. A multinomial regression model was built to investigate the association of baseline variables with the trajectory groups.
**PHQ-9 Trajectory Groups (Depression)**

- **Cluster 1** (n=62, 10.74%) High, nonresponse
- **Cluster 2** (n=49, 8.49%) High, response
- **Cluster 3** (n=137, 23.74%) Moderately severe, nonresponse
- **Cluster 4** (n=93, 16.12%) Moderately severe, response
- **Cluster 5** (n=106, 18.37%) Moderate, remission
- **Cluster 6** (n=130, 22.53%) Moderate, nonresponse

**GAD-7 Trajectory Groups (Anxiety)**

- **Cluster 1** (n=106, 18.37%) High, nonresponse
- **Cluster 2** (n=100, 17.33%) High, response
- **Cluster 3** (n=97, 16.81%) Moderate, nonresponse
- **Cluster 4** (n=115, 19.93%) Moderate, nonresponse
- **Cluster 5** (n=66, 11.44%) High, remission
- **Cluster 6** (n=93, 16.12%) Moderate, remission
Figure 2. The trajectory groups derived from PHQ-9 and GAD-7 scores
Figure 3. The overlap of patients between the PHQ-9 and GAD-7 trajectory groups
Supplementary information to:

Title: Longitudinal Trajectories of Symptom Change During Antidepressant Treatment Among Managed Care Patients with Co-Occurring Depression and Anxiety

1. Characteristics of the overall sample
   Across the entire sample, 42.98% of patients were characterized as responders for depressive symptoms, while 57.01% were non-responders. Baseline characteristics that significantly differed between groups included PHQ-9 and C-SSRS scores. Similar patterns emerged for anxiety symptoms, with 44.89% of patients in the responder groups, and 55.11% non-responders. Significant differences between groups were identified in the GAD-7 and PSQI scores. The types of antidepressants (SSRIs, SNRIs, TCAs, MAOIs, and atypical antidepressants) did not significantly differ across groups.

2. Baseline characteristics
   The baseline was defined as the first session and encounter where care was established at the outpatient behavioral health clinic. Demographic variables extracted from this session included age, gender, and self-reported race. Comorbidities included post-traumatic stress disorder (PTSD), bipolar disorder, and “suicidal behavior,” which comprised ICD-coded suicide attempt, suicidal ideation, and self-harm. The individual’s PHQ-9 and GAD-7 scores at baseline were also extracted. Several other prognostic factors were measured at baseline using multiple validated scoring instruments and questionnaires. These included the Alcohol Use Disorders Identification Test (AUDIT) which measures alcohol use and drinking behaviors; the Columbia-Suicide Severity Rating Scale (C-SSRS) measuring suicide risk; the Pittsburgh Sleep Quality Index (PSQI) which assesses sleep patterns and quality; the Patient-Reported Outcomes Measurement Information System (PROMIS) which evaluates overall physical, mental, and social wellbeing; and the Sheehan Disability Scale (SDS) which measures functional impairment in work/school, social, and family life. Lower symptom severity is indicated by lower scores on the AUDIT, C-SSRS, PSQI, and SDS questionnaires and a higher score on the PROMIS. The associations of these baseline characteristics with trajectories of depression and anxiety were investigated. All instruments were administered to the clinic patients as part of their routine clinical care.

3. Statistical Analysis
   GBTM was employed to stratify members of the cohort into clinically meaningful subgroups with statistically similar trajectories of their depression and anxiety symptoms. GBTM has been applied in clinical contexts to reveal patient groups with similar characteristics. Analysis was conducted in R (version 4.0.3) and GBTM was performed using the package lcmm. The number of groups and the degree of the polynomial in each of the trajectory groups were determined by using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to measure model fit with lower AIC and BIC values indicating better fit.

   After identifying trajectory groups using GBTM, further univariate and multivariate tests were performed to obtain associations between baseline prognostic factors and trajectory groups. In particular, the continuous characteristics (PHQ-9, GAD-7, AUDIT, C-SSRS,
PROMIS, PSQI, and SDS) were described using the mean and standard deviation (SD), while categorical characteristics (e.g., race and gender) were described using frequency counts. For the univariate analyses, we applied a student's t-test, Mann-Whitney U test, Chi-square test, and Fisher’s exact test where appropriate. Analysis of covariance (ANCOVA) for the between-trajectory group comparisons was also employed on the basis of the general linear model (GLM), adjusting for age at baseline. For the multivariate analyses, multinomial logistic regression models were used to assess associations of baseline prognostic factors with trajectory group membership, including their ability to predict membership based on pre-treatment patient characteristics.

4. Characteristics of the derived trajectory groups on anxiety symptom change

There were also six trajectory groups reflecting pathways of anxiety symptom change, which included high, nonresponse (Cluster 1), high, response (Cluster 2), moderate, nonresponse-1 (Cluster 3), moderate, nonresponse-2 (Cluster 4), high, remission (Cluster 5), and moderate, remission (Cluster 6). Patients in the trajectory group high, nonresponse underwent limited change during antidepressant treatment. Patients in the high, response group demonstrated a substantial decrease in anxiety symptoms. Patients in the high, remission group had severe anxiety at baseline but demonstrated the greatest symptom improvement and ultimately achieved levels below typical remission criteria (GAD-7 ≤ 5). Patients in the moderate, remission group initiated care with moderate anxiety symptoms and displayed steady improvement, also reaching levels below typical remission criteria. Although patients in high, remission and moderate, remission groups achieved the largest symptom improvements, the slope of high, remission group was steeper than that of moderate, remission group, which suggests that the speed of improvement was higher in high, remission group. We observed unremarkable improvement for patients in GAD-7 trajectory groups moderate, nonresponse-1 and moderate, nonresponse-2 groups. Patient characteristics for the GAD-7 trajectory groups are shown in eTable 1.

5. The associations of baseline patient characteristics with GAD-7 trajectory group membership

In terms of the GAD-7 trajectories, relative to being in GAD-7 trajectory Cluster 5 (High, remission), the probability of being in Cluster 1 (High, nonresponse) was higher and was associated with poorer overall health (lower PROMIS scores); the probability of being in GAD-7 trajectory Cluster 2 (High, response) was higher and was associated with more severity of anxiety prior to treatment (higher baseline GAD-7 scores). The probability of being in GAD-7 trajectory Cluster 3 (Moderate, nonresponse-1) was higher and was associated with less severity of anxiety prior to treatment (lower baseline GAD-7 scores); the probability was also higher if patients’ race was non-white. The probability of being in GAD-7 trajectory Cluster 4 (Moderate, nonresponse-2) was higher and was associated with less severity of anxiety prior to treatment and poorer sleep quality (lower baseline GAD-7 scores and higher PSQI); the probability was also higher if patients’ race was non-white. Finally, the probability of being in GAD-7 trajectory Cluster 6 (Moderate, remission) was higher and was associated with older, less severity of anxiety prior to treatment, and better overall health (higher age, PROMIS, and lower baseline GAD-7 scores).
6. eFigure 1: The inclusion/exclusion cascade for identifying study participants

**PHQ-9 Dataset**
(June 2016 – Sept. 2020)
\[N = 3,730\]

**GAD-7 Dataset**
(June 2016 – Sept. 2020)
\[N = 3,704\]

- Patients with both PHQ-9 and GAD-7
  \[n = 3,696\]
  - Excluded: Patients not prescribed an antidepressant
    \[n = 1,324\]
  - Excluded: Patients with <3 measurements during weeks 0-14
    \[n = 1,069\]
  - Excluded: Patients with baseline PHQ-9 <10 or GAD-7 <8
    \[n = 726\]

- Patients prescribed an antidepressant
  \[n = 2,372\]
  - Excluded: Patients with only PHQ-9 or GAD-7
    \[n = 42\]

- Patients with ≥3 measurements during weeks 0-14
  \[n = 1,303\]
  - Excluded: Patients with <3 measurements during weeks 0-14
    \[n = 1,069\]
  - Excluded: Patients with baseline PHQ-9 <10 or GAD-7 <8
    \[n = 726\]

- Final cohort for analysis
  \[n = 577\]
## eTable 1: Patient characteristics for the GAD-7 trajectory groups

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All patients</th>
<th>Cluster 1 (High, nonresponse)</th>
<th>Cluster 2 (High, response)</th>
<th>Cluster 3 (Moderate, nonresponse1)</th>
<th>Cluster 4 (Moderate, nonresponse2)</th>
<th>Cluster 5 (High, remission)</th>
<th>Cluster 6 (Moderate, remission)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>39.34 (12.87)</td>
<td>42.02 (12.9)</td>
<td>37.05 (12.36)</td>
<td>40.35 (13.5)</td>
<td>37.65 (13.75)</td>
<td>37.41 (10.67)</td>
<td>41.14 (12.03)</td>
<td>0.006595144</td>
</tr>
<tr>
<td>PHQ-9, mean (SD)</td>
<td>16.19 (4.26)</td>
<td>18.43 (4.34)</td>
<td>16.98 (4.31)</td>
<td>15.52 (4.08)</td>
<td>14.81 (3.93)</td>
<td>15.95 (3.92)</td>
<td>15.35 (3.72)</td>
<td>2.72E-09</td>
</tr>
<tr>
<td>GAD-7, mean (SD)</td>
<td>15.35 (3.81)</td>
<td>18.51 (2.19)</td>
<td>17.96 (2.01)</td>
<td>14.92 (2.8)</td>
<td>12.14 (2.63)</td>
<td>17.64 (2.45)</td>
<td>11.72 (3.29)</td>
<td>3.87E-65</td>
</tr>
<tr>
<td>AUDIT, mean (SD)</td>
<td>3.69 (1.61)</td>
<td>3.71 (1.65)</td>
<td>3.58 (1.76)</td>
<td>3.94 (1.58)</td>
<td>3.46 (1.49)</td>
<td>3.94 (1.72)</td>
<td>3.65 (1.38)</td>
<td>0.371474188</td>
</tr>
<tr>
<td>C-SSRS, mean (SD)</td>
<td>1.47 (1.2)</td>
<td>1.69 (1.32)</td>
<td>1.5 (1.16)</td>
<td>1.46 (1.19)</td>
<td>1.34 (1.19)</td>
<td>1.45 (1.02)</td>
<td>1.4 (1.2)</td>
<td>0.253010892</td>
</tr>
<tr>
<td>PROMIS, mean (SD)</td>
<td>13.57 (1.89)</td>
<td>12.73 (1.85)</td>
<td>13.54 (1.93)</td>
<td>13.66 (1.73)</td>
<td>13.84 (1.68)</td>
<td>13.59 (2.23)</td>
<td>14.1 (1.69)</td>
<td>9.27E-07</td>
</tr>
<tr>
<td>PSQI, mean (SD)</td>
<td>10.9 (2.7)</td>
<td>11.92 (2.61)</td>
<td>10.52 (2.88)</td>
<td>11.14 (2.44)</td>
<td>10.57 (2.89)</td>
<td>10.8 (2.28)</td>
<td>10.37 (2.58)</td>
<td>1.11E-05</td>
</tr>
<tr>
<td>SDS, mean (SD)</td>
<td>19.56 (4.72)</td>
<td>21.48 (5.11)</td>
<td>20.35 (4.86)</td>
<td>19.13 (4.64)</td>
<td>18.29 (3.94)</td>
<td>20.36 (4.17)</td>
<td>18.0 (4.34)</td>
<td>8.43E-15</td>
</tr>
<tr>
<td>Gender – Female (%)</td>
<td>373 (64.64)</td>
<td>70 (66.04)</td>
<td>63 (63.00)</td>
<td>64 (65.98)</td>
<td>72 (62.61)</td>
<td>42 (63.64)</td>
<td>62 (66.67)</td>
<td>0.983607073</td>
</tr>
<tr>
<td>Race – White (%)</td>
<td>313 (54.25)</td>
<td>58 (54.72)</td>
<td>52 (52.00)</td>
<td>50 (51.55)</td>
<td>57 (49.57)</td>
<td>42 (63.64)</td>
<td>54 (58.06)</td>
<td>0.492842977</td>
</tr>
</tbody>
</table>
8. eTable 2: Associations of baseline patient characteristics with PHQ-9 trajectory group membership

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>GAD-7 Class 1 (High, nonresponse): RR (95%CI) &amp; p-value</th>
<th>GAD-7 Class 2 (High, response): RR (95%CI) &amp; p-value</th>
<th>GAD-7 Class 3 (Moderate, nonresponse1): RR (95%CI) &amp; p-value</th>
<th>GAD-7 Class 4 (Moderate, nonresponse2): RR (95%CI) &amp; p-value</th>
<th>GAD-7 Class 5 (High, remission): RR (95%CI) &amp; p-value</th>
<th>GAD-7 Class 6 (Moderate, remission): RR (95%CI) &amp; p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.01 (0.98-1.04), p = .402</td>
<td>1.0 (0.97-1.02), p = .763</td>
<td>1.02 (1.0-1.05), p = .106</td>
<td>1.01 (0.98-1.04), p = .361</td>
<td>Reference</td>
<td>1.04 (1.01-1.07), p = .012</td>
</tr>
<tr>
<td>Gender – Female</td>
<td>0.88 (0.45-1.72), p = .715</td>
<td>0.93 (0.48-1.81), p = .838</td>
<td>1.17 (0.58-2.36), p = .661</td>
<td>0.96 (0.46-2.01), p = .911</td>
<td>Reference</td>
<td>1.1 (0.5-2.42), p = .806</td>
</tr>
<tr>
<td>Race – Non-White</td>
<td>1.28 (0.66-2.47), p = .47</td>
<td>1.6 (0.83-3.07), p = .16</td>
<td>2.05 (1.03-4.08), p = .04</td>
<td>2.47 (1.18-5.16), p = .016</td>
<td>Reference</td>
<td>1.99 (0.92-4.32), p = .082</td>
</tr>
<tr>
<td>AUDIT</td>
<td>0.95 (0.79-1.15), p = .621</td>
<td>0.85 (0.69-1.04), p = .105</td>
<td>1.06 (0.87-1.29), p = .588</td>
<td>0.88 (0.7-1.1), p = .261</td>
<td>Reference</td>
<td>0.91 (0.72-1.16), p = .463</td>
</tr>
<tr>
<td>C-SSRS</td>
<td>0.84 (0.61-1.16), p = .282</td>
<td>0.88 (0.63-1.24), p = .471</td>
<td>0.91 (0.65-1.28), p = .603</td>
<td>0.76 (0.52-1.11), p = .162</td>
<td>Reference</td>
<td>0.79 (0.52-1.19), p = .256</td>
</tr>
<tr>
<td>LEC-5</td>
<td>1.03 (0.95-1.13), p = .47</td>
<td>1.06 (0.98-1.16), p = .163</td>
<td>0.99 (0.9-1.08), p = .749</td>
<td>1.0 (0.91-1.1), p = .977</td>
<td>Reference</td>
<td>0.97 (0.87-1.08), p = .597</td>
</tr>
<tr>
<td>PCL-C</td>
<td>1.1 (1.0-1.21), p = .053</td>
<td>0.99 (0.9-1.08), p = .784</td>
<td>1.07 (0.97-1.18), p = .196</td>
<td>1.1 (0.99-1.22), p = .09</td>
<td>Reference</td>
<td>1.03 (0.91-1.15), p = .664</td>
</tr>
<tr>
<td>PROMIS</td>
<td>0.82 (0.72-0.94), p &lt; .001</td>
<td>1.01 (0.89-1.15), p = .895</td>
<td>1.19 (1.05-1.36), p &lt; .001</td>
<td>1.5 (1.31-1.73), p &lt; .001</td>
<td>Reference</td>
<td>1.53 (1.33-1.77), p &lt; .001</td>
</tr>
<tr>
<td>PSQI</td>
<td>1.0 (0.87-1.16), p = .964</td>
<td>0.88 (0.77-1.02), p = .081</td>
<td>1.2 (1.04-1.39), p = .016</td>
<td>1.21 (1.03-1.42), p = .019</td>
<td>Reference</td>
<td>1.15 (0.97-1.37), p = .1</td>
</tr>
<tr>
<td>SDS</td>
<td>0.94 (0.87-1.02), p = .121</td>
<td>0.98 (0.91-1.05), p = .517</td>
<td>0.98 (0.91-1.07), p = .7</td>
<td>1.02 (0.94-1.12), p = 6</td>
<td>Reference</td>
<td>1.02 (0.93-1.12), p = .72</td>
</tr>
<tr>
<td>PHQ-9</td>
<td>1.1 (1.0-1.22), p = .054</td>
<td>1.1 (1.0-1.22), p = .05</td>
<td>1.06 (0.95-1.17), p = .302</td>
<td>1.13 (1.01-1.26), p = .036</td>
<td>Reference</td>
<td>1.19 (1.05-1.34), p &lt; .001</td>
</tr>
<tr>
<td>GAD-7</td>
<td>1.03 (0.89-1.2), p = .692</td>
<td>1.06 (0.92-1.23), p = .424</td>
<td>0.65 (0.56-0.76), p &lt; .001</td>
<td>0.48 (0.41-0.57), p &lt; .001</td>
<td>Reference</td>
<td>0.46 (0.39-0.55), p &lt; .001</td>
</tr>
</tbody>
</table>

9. eTable 3: The process for choosing the best number of groups based on PHQ-9 score in GBTM with AIC and BIC

<table>
<thead>
<tr>
<th>Number of Groups</th>
<th>Akaike Information Criterion (AIC)</th>
<th>Bayesian Information Criterion (BIC)</th>
<th>Class distribution (# and %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>48077.37</td>
<td>48125.46</td>
<td>class1 class2 N 206 379 % 35.21 64.79</td>
</tr>
<tr>
<td>3</td>
<td>47098.54</td>
<td>47159.74</td>
<td>class1 class2 class3 N 151 264 170 % 25.81 45.13 29.06</td>
</tr>
<tr>
<td>4</td>
<td>46660.38</td>
<td>46734.69</td>
<td>class1 class2 class3 class4 N 56 147 254 128 % 9.57 25.13 43.42 21.88</td>
</tr>
<tr>
<td>5</td>
<td>46424.13</td>
<td>46611.56</td>
<td>class1 class2 class3 class4 class5 N 59 103 150 114 159 % 10.09 17.61 25.64 19.49 27.18</td>
</tr>
<tr>
<td>6</td>
<td>46278.78</td>
<td>46379.33</td>
<td>class1 class2 class3 class4 class5 class6 N 45 111 109 133 113 74 % 7.69 18.97 18.63 22.74 19.32 12.65</td>
</tr>
<tr>
<td>7</td>
<td>46280.94</td>
<td>46394.6</td>
<td>class1 class2 class3 class4 class5 class6 class7 N 36 95 145 139 75 95 0 % 6.15 16.24 24.79 23.76 12.82 16.24 0</td>
</tr>
</tbody>
</table>
10. *eTable* 4: The process for choosing the best number of groups based on GAD-7 score in GBTM with AIC and BIC

<table>
<thead>
<tr>
<th>Number of Groups</th>
<th>Akaike Information Criterion (AIC)</th>
<th>Bayesian Information Criterion (BIC)</th>
<th>Class distribution (# and %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>47313.76</td>
<td>47339.99</td>
<td>class1 class2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N 276 309</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% 47.18 52.82</td>
</tr>
<tr>
<td>3</td>
<td>46515.1</td>
<td>46554.45</td>
<td>class1 class2 class3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N 132 248 205</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% 22.56 42.39 35.04</td>
</tr>
<tr>
<td>4</td>
<td>46227.12</td>
<td>46279.58</td>
<td>class1 class2 class3 class4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N 87 193 201 104</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% 14.87 32.99 34.36 17.78</td>
</tr>
<tr>
<td>5</td>
<td>46006.89</td>
<td>46072.47</td>
<td>class1 class2 class3 class4 class5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N 90 127 161 114 93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% 15.38 21.71 27.52 19.49 15.9</td>
</tr>
<tr>
<td>6</td>
<td>45908.99</td>
<td>45987.68</td>
<td>class1 class2 class3 class4 class5 class6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N 87 118 97 106 100 77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% 14.87 20.17 16.58 18.12 17.09 13.16</td>
</tr>
<tr>
<td>7</td>
<td>45914.99</td>
<td>46006.8</td>
<td>class1 class2 class3 class4 class5 class6 class7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N 87 118 97 106 100 77 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% 14.87 20.17 16.58 18.12 17.09 13.16 0</td>
</tr>
</tbody>
</table>