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

Rudolf Uher, PhD; Ole Mors, PhD; Marcella Rietschel, MD; Aleksandra Rajewska-Rager, MD; Ana Petrovic, PhD; Astrid Zobel, MD; Neven Henigsberg, MD; Julien Mendlewicz, MD, PhD; Katherine J. Aitchison, PhD; Anne Farmer, MD; and Peter McGuffin, PhD

ABSTRACT

Objective: The timing and rate of improvement after the initiation of an antidepressant has implications for establishing the mechanism of antidepressant action and for answering the clinically relevant question of how long an appropriate trial of antidepressant medication should be. We explore the individual trajectories of relative change in depression severity to establish what proportion of individuals experience early and late onset of improvement.

Method: Longitudinal latent class analysis was applied in a secondary analysis of data obtained from the Genome-Based Therapeutic Drugs for Depression (GENDEP) study. In the GENDEP trial, conducted in 9 European academic psychiatry centers from July 2004 to June 2008, 811 treatment-seeking adult subjects with *DSM-IV* major depression received escitalopram or nortriptyline for 12 weeks. Montgomery-Asberg Depression Rating Scale measurements were taken weekly. The secondary analysis reported in this article was conducted in 2010.

Results: A model with 9 latent classes provided a good description of the individual trajectories of symptom change over time. These classes included 3 nonresponder classes, 3 classes with varying degrees of improvement concentrated in the first 3 weeks (early improvement), and 3 classes with varying degrees of improvement that was more prominent in the second 3 weeks than in the first 3 weeks (delayed improvement). More than half of the subjects who eventually reached remission showed a pattern of delayed improvement, and their eventual outcome could not be predicted from early time points. Early marked response occurred more frequently in subjects treated with nortriptyline than in those treated with escitalopram (12.9% vs 7.5%, $\chi^2 = 6.29$, $P = .01$). Delayed complete remission occurred more frequently in subjects treated with escitalopram than in those treated with nortriptyline (13.6% vs 6.1%, $\chi^2 = 11.52$, $P = .0007$).

Conclusions: Both early and delayed improvement are common. Although early changes are maintained, the eventual outcome of 12-week antidepressant treatment can be accurately predicted only after 8 weeks.

Trial Registration: http://www.controlled-trials.com Identifier: ISRCTN03693000.

Antidepressant medication remains the mainstay of treatment for major depression. However, a large proportion of individuals with depression do not respond to the first antidepressant, and they require switching to another antidepressant or augmentation with a second agent. An important question is how long an antidepressant should be continued before an attempt is declared failed and another medication or augmentation is attempted. Treatment with an antidepressant that is not effective for a given individual should not be longer than is necessary to establish lack of efficacy, as prolonged ineffective treatment is associated with extended suffering, disability, and risk of suicide.[1,2] If, on the other hand, a treatment that has the potential to be effective is curtailed prematurely, the therapeutic trial is wasted, with subsequent narrowing of therapeutic repertoire and exposure to less well-tolerated antidepressants or unnecessary drug combinations. Multiple prematurely terminated treatment trials can also prolong time to eventual recovery, cause frustration, and increase the risk of suicide.

Professional opinion on the length of an adequate antidepressant trial has changed substantially in the last decade. Traditionally, it was believed that therapeutic response to antidepressants occurs after a delay of 2 to 3 weeks and cannot be fully evaluated until after 8 to 12 weeks.[3] Therefore, patients were informed not to expect improvement until after several weeks of treatment, and early changes were considered to be nonspecific and possibly transitory.[3,4] Accordingly, most treatment guidelines recommend 4 to 6 weeks as a minimum duration of an antidepressant treatment trial.[5–7] This view has been challenged by new evidence demonstrating that differences between an active antidepressant and placebo emerge within the first week of treatment[8–13] and that early improvement predicts sustained response.[10,11,14,15] It has been suggested that lack of early improvement of at least 20% over the first 2 weeks of treatment may constitute a reason for medication change.[11] Such a major change in treatment guidelines requires replication and careful scrutiny of evidence, which are among the aims of the present study.

A synthesis of evidence shows that antidepressants separate from placebo within the first week of treatment and that early improvement is usually maintained.[9,12] However, it is also clear that the therapeutic effect of antidepressants continues to grow for at least 6 weeks of treatment.[12] Since data from clinical trials are usually presented as means of large groups, these findings are consistent with either continuous improvement with early onset

FOR CLINICAL USE

◆ The time course of improvement after initiation of an antidepressant varies from patient to patient. While some patients experience rapid improvement in the first 2 to 3 weeks following initiation of an antidepressant, others start significantly improving only after a delay of 2 to 4 weeks. Early and delayed onsets of improvement are approximately equally common.

◆ If tolerated, an antidepressant should be continued for 6 to 8 weeks in adequate doses, even if there is no discernible therapeutic effect in the first 2 weeks of treatment. Patients who do not achieve improvement in the first 2 weeks of treatment still have a 40% to 50% chance of significant improvement if they persevere with treatment for 6 to 8 weeks.

◆ Six to eight weeks of treatment constitutes an adequate trial of an antidepressant. If significant improvement is not achieved after 6 to 8 weeks of treatment with adequate doses, the clinician should consider changing treatment.

in all responders or with a mixture of individuals' experiencing early and late improvement. The distinction between these 2 possibilities has implications for individualized decisions about continuing or discontinuing antidepressants. If the observed pattern is due to a mixture of early and late responders but the means are generalized to all individuals with depression, those with a delayed pattern of response will have their treatment curtailed prematurely and miss the opportunity to achieve remission.

Individual differences in the trajectory of response can be studied with latent class analyses separating groups of individuals who follow distinct trajectories.[16,17] One commonly applied type of latent variable analysis is the growth mixture model (GMM), which provides efficient description of longitudinal data and may improve the analysis of clinical trials.[18–23] However, GMM restricts the identifiable classes to those that follow smooth curves and does not allow trajectories with different onset of improvement. Application of a more general longitudinal latent class analysis, which allows all possible shapes of latent trajectories, may maximize the description of individual differences and uncover more qualitatively different trajectories of change, such as early and late responders. In the present study, we probe the onset of antidepressant response and the prediction of final outcome from initial response using both longitudinal latent class analysis and a traditional approach.

METHOD

Study Design and Sample

The analyses were based on 811 treatment-seeking adults with a depressive episode fulfilling *DSM-IV* criteria[24] for a major depressive episode and/or *ICD-10* criteria[24] for a depressive episode of at least moderate severity. Participants were treated with escitalopram or nortriptyline in 9 European academic psychiatry centers from July 2004 to June 2008 as part of the Genome-based Therapeutic Drugs for Depression (GENDEP) study, a partially randomized, multicenter, open-label study. The sample and primary outcomes have been described elsewhere.[25] (The secondary analysis reported in this article was conducted in 2010.) The original study protocol was approved by the research ethics boards of all participating centers, and written informed consent was obtained from all participants. The study was registered at http://www.controlled-trials.com (Identifier: ISRCTN03693000). The mean age of participants was 43 years (SD = 12 years). Participants with no contraindications were randomly allocated to receive flexible-dosage nortriptyline (n = 235) or escitalopram (n = 233) for 12 weeks. Patients with contraindications for one of the drugs were allocated nonrandomly to the other antidepressant, 225 to escitalopram and 118 to nortriptyline. Escitalopram was initiated at 10 mg daily and increased to 15 mg daily within the first 2 weeks unless adverse effects limited dose increase, and it could be further increased up to 30 mg (median = 15 mg, mean = 17 mg). Nortriptyline was initiated at 50 mg daily and titrated to 100 mg daily within the first 2 weeks unless adverse effects limited dose increase, and it could be further increased up to 200 mg (median = 100 mg, mean = 107 mg). The time course of change in depressive symptoms was modeled in 806 individuals with a valid baseline and at least 1 postbaseline measurement on the originally allocated medication. Of these, 628 (78%) completed 8 weeks, and 527 (6.5%) completed 12 weeks taking the allocated antidepressant. The weekly measurements of depression severity before study exit (completion or dropout) were 93% complete.

Measures of Depression Severity

The primary outcome measure was the Montgomery-Asberg Depression Rating Scale (MADRS),[26] administered at baseline and then in weekly intervals over the 12 weeks of treatment by psychologists and psychiatrists with interrater reliability of 0.9.[27] At baseline, the participants had a mean MADRS score of 28.7 (SD = 6.7). Response was defined as a 50% reduction in MADRS score from baseline.[25] Remission was defined as a MADRS score of 10 or less at week 12, with missing week-12 data estimated as the best unbiased linear estimate from mixed-effect models.[25]

Longitudinal Latent Class Analysis

To minimize the influence of baseline severity and avoid classes defined by degrees of severity across the study rather than the course of change, depression severity at

Figure 1. The 9 Latent Trajectories Model[a]

Legend:
◇ Class 1, 2% (n = 11)
○ Class 2, 8% (n = 64)
□ Class 3, 10% (n = 81)
▲ Class 4, 17% (n = 140)
◆ Class 5, 14% (n = 114)
■ Class 6, 10% (n = 79)
△ Class 7, 18% (n = 148)
△ Class 8, 10% (n = 84)
● Class 9, 11% (n = 85)

Axis labels: Baseline Severity of Depression, % (y-axis); Study Week (x-axis)
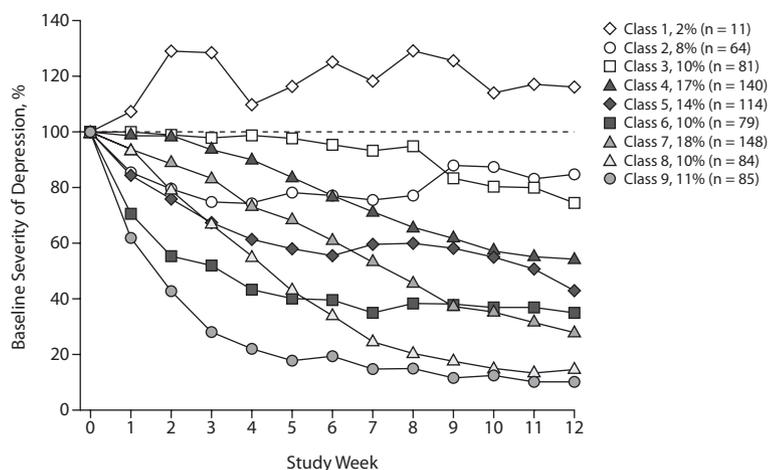
[a]Model estimates of class means are plotted, which for longitudinal latent class analysis are equal to observed means of individuals belonging to each latent class. Latent trajectory classes are ordered according to the relative severity of depression at study endpoint.

each week was expressed as a percentage of the baseline MADRS score for each participant. Longitudinal latent class analysis (LLCA) was fitted using maximum likelihood estimation in Mplus, version 5.21 (Muthén & Muthén, Los Angeles, California).[28] Longitudinal latent class analysis maximizes the variability in MADRS that is explained by the latent classes or prototype trajectories of change.[17,29] Through the assumption of conditional independence, all variance is explained by class membership alone. All possible shapes of trajectories are allowed. We applied a series of LLCA models with increasing numbers of classes and compared models using the Bayesian Information Criterion (BIC), with a smaller BIC indicating a better model.[30,31] In addition, we performed the bootstrapped likelihood-ratio test to confirm that a specific model fit data significantly better than a model with 1 fewer class.[32] Another criterion was the number of subjects in the smaller classes, as classes with only 1 participant or a few do not add advantage in providing a summary measure. Quality of classification was assessed with entropy (ranging from 0 = random to 1 = perfect) with values of 0.8 or more indicating a clear separation of classes.[33] As it was our aim to test the classification of course, information on drug was initially not included in the model.

Other Statistical Analyses

The influence of categorical variables (eg, drug) on class membership was tested using the Pearson χ^2 test and logistic regression. Differences in continuous variables (eg, depression severity) were tested using linear regression models. Agreement between 2 categorical variables was quantified by the Cohen κ coefficient. For prediction of latent class membership, weighted kappa values with quadratic weights according to the order of classes at week 12 were used in addition to absolute agreement of each class membership. For agreement between 2 dichotomous variables, such as

the prediction of endpoint response (≥ 50% reduction at week 12) from a dichotomous variable of early improvement (≥ 20% reduction at week 2) as previously reported by Szegedi et al,[10,11] sensitivity, specificity, and positive and negative predictive values were also calculated. Dose titration was compared using Cox proportional hazard regressions with time to reach a midrange dose, half-way between the lowest effective and highest recommended dose, ie, 15 mg for escitalopram and 100 mg for nortriptyline.[34] These analyses were implemented in STATA, release 10 (StataCorp LP, College Station, Texas).[35]

RESULTS

Longitudinal Latent Class Analysis

A series of LLCA models with increasing numbers of latent classes was fitted to describe the individual trajectories of symptom change among the 806 participants with valid postbaseline measurements. Models with up to 5 classes separated groups of participants with varying degrees of change. Models with 6 or more classes separated qualitatively distinct trajectories of participants who experienced improvement at different stages of the 12-week treatment period. Information criteria continued decreasing, indicating that models with more classes provide significantly better descriptions of the dataset than models with fewer classes. This was confirmed with bootstrapped likelihood-ratio tests (eg, a model with 9 classes fit significantly better than a model with 8 classes; $P < .0001$; eTable 1 [available at PSYCHIATRIST.COM]). However, in models with 10 or more classes, the additional classes contained a single participant with an unusual trajectory of change. Therefore, we selected the 9-class model for further analyses.

The LLCA 9-class model is shown in Figure 1 and described in Table 1. Class 1 groups the few participants who experienced worsening of depressive symptoms. Classes 2 and 3 group nonresponders with minimal improvement occurring either initially or toward the end of the study. The remaining 6 classes grouped participants with varying degrees of response occurring either in the initial 3 weeks or in the middle stages of the study. Classes 5, 6, and 9 comprised individuals who experienced various degrees of reduction in depressive symptoms (moderate, marked, or dramatic, respectively) over the first 3 weeks, with relatively little change thereafter. These *early* improver classes comprised 278 study participants (34%). In contrast, classes 4, 7, and 8 grouped participants with a delayed onset of improvement, starting in the third week and continuing up to week 8 or 9, with symptom reduction in weeks 4 to 6 greater than that occurring in weeks 1 to 3. These 3 *delayed* improver classes comprised 372 study participants (46%). Few participants experienced substantial changes in depression severity in

Table 1. Description of the 9 Latent Trajectories of Change in Symptoms of Antidepressant-Treated Subjects

Class[a]		n	Depression Severity (mean MADRS scores)			Improvement (MADRS scores), %			Subjects Reaching Favorable Endpoint (MADRS scores), %		Attrition,[d] %
			Week 0	Week 3	Week 12	Week 0–3	Week 3–6	Week 0–12	Response[b]	Remission[c]	
1	Worsening	11	22.6	28.0	28.2	−28.7	3.5	−16.3	0.0	9.1	18.2
2	Early minimal improvement	64	29.8	22.5	23.7	25.2	−2.2	15.1	0.0	3.1	23.4
3	Late minimal improvement	81	27.9	27.7	19.7	2.2	2.4	25.5	6.2	8.6	25.9
4	Delayed partial response	140	29.2	27.4	15.9	6.2	16.3	45.8	25.7	15.7	27.1
5	Early partial response	114	29.9	19.7	12.9	32.5	11.8	57.1	68.4	27.2	15.8
6	Early marked response	79	28.8	15.3	9.9	48.0	12.5	65.2	83.5	57.0	15.2
7	Delayed marked response	148	29.4	24.6	8.2	16.7	22.6	72.3	84.5	60.8	15.5
8	Delayed complete remission	84	27.6	18.8	3.9	33.0	33.2	85.6	97.6	90.5	10.7
9	Early complete remission	85	27.8	7.9	2.9	72.1	8.8	90.0	100.0	94.1	18.8

[a]Latent trajectory classes are ordered according to the relative severity of depression at study endpoint. [b]Defined as a 50% reduction in MADRS score from baseline. [c]Defined as a MADRS score ≤ 10 at week 12. [d]Denotes subjects who left the study before week 12.
Abbreviation: MADRS = Montgomery-Asberg Depression Rating Scale.

Table 2. Comparison of Trajectories of Change in Depressive Symptoms Between Participants Treated With Escitalopram and Nortriptyline[a]

Latent Trajectory Class		Entire Sample						Randomly Allocated Participants					
		Escitalopram (n = 464)		Nortriptyline (n = 342)		Comparison[b]		Escitalopram (n = 230)		Nortriptyline (n = 229)		Comparison[b]	
		n	%	n	%	χ^2	P	n	%	n	%	χ^2	P
1	Worsening	7	1.5	4	1.2	0.17	.6837	3	1.3	1	0.4	1.01	.3153
2	Early minimal improvement	33	7.1	31	9.1	1.03	.3096	20	8.7	22	9.6	0.10	.7470
3	Late minimal improvement	51	11.0	30	8.8	1.05	.3054	23	10.0	12	5.2	3.72	.0537
4	Delayed partial response	78	16.8	62	18.1	0.24	.6213	31	13.5	41	17.9	1.63	.2023
5	Early partial response	60	12.9	54	15.8	1.33	.2491	37	16.1	38	16.6	0.02	.9002
6	Early marked response	35	7.5	44	12.9	**6.29**	**.0121**	20	8.7	34	14.9	**4.08**	**.0435**
7	Delayed marked response	88	19.0	60	17.5	0.25	.6159	37	16.1	46	20.1	1.17	.2788
8	Delayed complete remission	63	13.6	21	6.1	**11.52**	**.0007**	34	14.8	13	5.7	**10.37**	**.0013**
9	Early complete remission	49	10.6	36	10.5	0.00	.9932	25	10.9	22	9.6	0.21	.6455
						19.77	**.0113**					**20.13**	**.0098**

[a]Association between antidepressant and each outcome trajectory class are compared using Pearson χ^2 with 1 degree of freedom. The last line in the table gives the results of an overall association between drug and class membership tested with a χ^2 test with 8 degrees of freedom. [b]Drug differences significant at $P < .05$ are highlighted in bold.

the last 3 weeks. Entropy of 0.82 reflected good classification, with most subjects clearly allocated to a single class and predicted class trajectories closely corresponding to observed individual trajectories (eFigure 1).

Influence of Baseline Severity, Drug, Sex, and Age

Inclusion of baseline severity, drug, random allocation, age, and sex as covariates did not change the LLCA results in terms of model fit, trajectory shape, or class membership. Compared to other classes, members of class 1 (Worsening) had lower scores of depression at baseline (linear regression: b = −6.28, 95% CI, −10.67 to −2.48; P = .0017). There were no differences in baseline severity between other classes (all P values > .05; Table 1). Drug was significantly related to class (Table 2). More nortriptyline-treated participants followed a trajectory of early marked response (class 6), and more escitalopram-treated individuals followed a trajectory of delayed complete remission (class 8). These differences remained in a sensitivity analysis restricted to randomly allocated participants (Table 2). Class membership was unrelated to sex ($\chi^2_8 = 6.93$, P = .5443). Age significantly predicted trajectory, with a younger age increasing the probability of subjects' following trajectories with better outcomes (classes 6 to 9; linear regression of age on class $F_{8,797} = 3.99$; P = .0001). Drug-age interactions were nonsignificant.

Antidepressant Dose Titration

To test whether trajectories of late and early improvement were influenced by the differential rate of dose titration, we explored the relationship between dose titration and latent class membership (eTable 2). Among escitalopram-treated subjects, the 2 classes with the worst outcomes (classes 1 and 2) had quicker titration, and they reached the midrange dose earlier than other classes (class 1 hazard ratio [HR] = 2.41, 95% CI, 1.09–5.32; P = .0302; class 2 HR = 1.90, 95% CI, 1.18–3.06; P = .0085). The class with the best outcome (class 9, early complete remission) had slower titration, and it reached the midrange dose later (Cox regression: HR = 0.47, 95% CI, 0.26–0.84; P = .0118). Other classes did not differ in dose titration, and there was no significant difference between the 3 classes with early improvement and the 3 cases with delayed improvement (Cox regression: HR = 0.84, 95% CI, 0.61–1.15; P = .2800). Among nortriptyline-treated subjects, there were no significant class differences in the rate of dose titration (all P values > .05).

DISCUSSION

Table 3. Prediction of Antidepressant Treatment Outcome From Earlier Time Points[a]

| Week | Prediction of Latent Class Membership[b] | | | Prediction of Response[c] | | Prediction of Remission[d] | |
| | Correctly Classified, % | κ | | Correctly Classified, % | κ | Correctly Classified, % | κ |
		Absolute	Weighted				
1	22.48	0.07	0.08	63.70	0.08	52.89	0.03
2	32.01	0.21	0.36	66.74	0.28	63.22	0.26
3	37.29	0.27	0.46	69.02	0.34	66.36	0.32
4	44.22	0.35	0.48	73.85	0.43	69.28	0.38
5	51.59	0.44	0.60	73.67	0.42	67.37	0.34
6	60.09	0.54	0.69	75.23	0.45	68.37	0.36
7	67.68	0.63	0.77	77.65	0.50	70.07	0.39
8	75.27	0.71	0.84	80.92	0.56	71.13	0.42
9	85.13	0.83	0.91	83.92	0.63	75.53	0.50
10	89.53	0.88	0.94	85.19	0.65	74.53	0.48
11	96.05	0.95	0.98	87.67	0.72	75.71	0.51
12	Reference	Reference	Reference	Reference	Reference	Reference	Reference

[a]κ Coefficient of agreement between classification of projected trajectory class, response, and remission based on percentage improvement at each week. [b]For latent trajectory class membership, weighted κ was calculated with order of classes according to endpoint severity and with quadratic weights. [c]Defined as a 50% reduction in MADRS score from baseline. [d]Defined as a MADRS score ≤ 10 at week 12.

Estimation of Final Outcome From Earlier Time Points

To address the question of how long a trial of an antidepressant is required to establish whether an individual is likely to achieve a satisfactory response, we used the continuous variable percentage improvement at weeks 1 to 11 to predict categorical outcomes at week 12 defined either as latent class membership or as response (50% improvement) or remission (endpoint MADRS score ≤ 10; Table 3). From data at week 2, only 32% of subjects were correctly classified as to their trajectory class, as subjects with delayed response could not be distinguished from those who were not going to improve. Similarly, 67% and 63% of subjects could be correctly classified as eventual responders and remitters, respectively, on the basis of week 2 data, a significant but modest increase above chance prediction, quantified by coefficients of agreement (κ = 0.28 for response and κ = 0.26 for remission). The percentage of subjects correctly classified continued to increase in a linear fashion as data from consecutive weeks were added. At week 8, the coefficients of agreement reached excellent levels (κ ≥ 0.8) for predicting class membership and acceptable levels for predicting response. Prediction of remission was less accurate, and it did not increase further after week 9.

The prediction of endpoint response (≥ 50% reduction at week 12) from a dichotomous variable of early improvement (≥ 20% reduction at week 2) correctly classified 65% of subjects, compared to a chance agreement of 59%, resulting in a κ coefficient of 0.30, sensitivity of 0.61, specificity of 0.70, positive predictive value of 0.75, and negative predictive value of 0.56.

Individual Outliers

Three subjects set apart as separate classes in LLCA models with 10 to 12 classes were examined because their symptoms followed unusual trajectories. In each case, the unusual symptom change trajectory could be explained as a reaction to dose change or a severe stressful life event (eFigure 2).

The analysis of individual trajectories of change demonstrates that both early and delayed improvement are common during treatment with SSRIs and tricyclic antidepressants. Among those who eventually show good response to a 12-week course of antidepressant treatment, 51% show delayed response, which cannot be predicted from measurements in the first 2 weeks. The delayed onset of improvement is not a function of slower dose titration, baseline severity, sex, or age.

Dose titration did not differ between trajectories with early and delayed onset of improvement. Subjects with very little or no improvement tended to receive higher doses of antidepressants earlier. These data are compatible with a scenario in which clinicians increase dosage in the absence of improvement rather than one in which higher doses lead to better or quicker improvement. This suggests that the determinants of symptom change trajectories are characteristics of subjects rather than treatment regimen.

Nine prototype trajectories that differ in the rate and degree of early and later improvement were needed to adequately describe the time course of change in depression severity in most individuals. This is consistent with the results of a meta-analysis in which improvement started in the first week of antidepressant treatment and continued for at least 6 weeks.[12] The present study suggests that this pattern reflects the averaging across a mixture of individuals with early and delayed improvement. This heterogeneity means that early changes are more informative for some patients than for others. Early improvement is usually maintained. However, lack of early improvement does not consistently predict lack of later response (negative predictive value of only 0.56) and therefore cannot constitute a basis for early termination of an antidepressant trial. The lack of clinical or demographic predictors of delayed improvement means that early prediction of treatment response is not plausible unless new markers predicting sustained response are established. Some promising advances have already been made.[36]

Of the 416 GENDEP participants who did not achieve 20% reduction in symptoms after 2 weeks of treatment, 267(64.2%) showed various degrees of delayed response, and 185 (44.5%) experienced 50% or greater reduction in depressive symptoms by week 12. These results strongly indicate that discontinuing an antidepressant after 2 weeks of treatment on the basis of the absence of an early improvement of 20% would be premature. The accuracy of final outcome prediction increases linearly with each consecutive

week, suggesting no obvious decision point. Therefore, we suggest that the decision of how long an antidepressant should be continued should take into account all available information, including personal and family history of response to antidepressant medication and the advantages and risks of medication change.

We did not replicate the reported high sensitivity and specificity of the prediction from early improvement at week 2 to response at endpoint.[10,11] The negative predictive value, reflecting the proportion of eventual nonresponse among subjects who do not show an early improvement of 20% or more after 2 weeks of treatment, was substantially lower in our sample (0.47 to 0.56) than the values reported by Szegedi and colleagues (0.82 to 1.00).[10,11] Different choice of antidepressants, titration regiments, inclusion criteria, and placebo control may explain the difference. The use of the last-observation-carried-forward procedure to replace missing data in the analyses by Szegedi and colleagues may also have contributed. If missing later values are replaced by earlier measurements, the correspondence between earlier and later time points increases spuriously. It has been repeatedly demonstrated that the last-observation-carried-forward procedure leads to significant bias.[25,37,38]

The present study exemplifies the use of longitudinal latent class analysis for descriptive and classification purposes. This method has inherent strengths and limitations. LLCA requires a large number of parameters to describe the data and is less parsimonious than the more commonly used GMM.[18] However, LLCA has the advantage that all possible shapes of trajectories are allowed and information on severity at all points in the trial is concentrated in the classifier variable, resulting in close correspondence between class means and individual observed values (eFigure 1). However, LLCA is computationally demanding and may not be routinely applicable. Without complex statistics, a useful approximation of the delayed response is the proportion of change in depression severity in weeks 4 to 6 to change in weeks 1 to 3. A ratio larger than 1 corresponds to a delayed onset of response.

When interpreting the present results, it is important to consider that GENDEP is a partially randomized study. The inclusion of subjects who could not be randomly allocated increases the generalizability of results to clinical populations but complicates the interpretation of between-drug comparisons. The classification of individuals into latent trajectories did not differ between randomly and nonrandomly allocated individuals, and all drug differences were confirmed in sensitivity analyses restricted to randomly allocated individuals.

In conclusion, the present study demonstrates substantial individual differences in the time-course of therapeutic change during antidepressant treatment, with both early and delayed response commonly occurring. The low predictive power of early changes for eventual outcome cautions against early discontinuation of antidepressants in the absence of initial effect. Future research should focus on the identification of novel predictors of outcome.

REFERENCES

1. Perroud N, Uher R, Marusic A, et al. Suicidal ideation during treatment of depression with escitalopram and nortriptyline in genome-based therapeutic drugs for depression (GENDEP): a clinical trial. *BMC Med.* 2009;7(1):60.

2. Zisook S, Trivedi MH, Warden D, et al. Clinical correlates of the worsening or emergence of suicidal ideation during SSRI treatment of depression: an examination of citalopram in the STAR*D study. *J Affect Disord.* 2009;117(1–2):63–73.

3. Quitkin FM, Rabkin JG, Ross D, et al. Identification of true drug response to antidepressants: use of pattern analysis. *Arch Gen Psychiatry.* 1984;41(8):782–786.

4. Quitkin FM, McGrath PJ, Stewart JW, et al. Can the effects of antidepressants be observed in the first two weeks of treatment? *Neuropsychopharmacology.* 1996;15(4):390–394.

5. Bauer M, Bschor T, Pfennig A, et al; WFSBP Task Force on Unipolar Depressive Disorders. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders in Primary Care. *World J Biol Psychiatry.* 2007; 8(2):67–104.

6. Anderson IM, Ferrier IN, Baldwin RC, et al. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2000 British Association for Psychopharmacology guidelines. *J Psychopharmacol.* 2008;22(4):343–396.

7. Lam RW, Kennedy SH, Grigoriadis S, et al; Canadian Network for Mood and Anxiety Treatments (CANMAT). Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults, 3: pharmacotherapy. *J Affect Disord*. 2009;117(suppl 1):S26–S43.

8. Katz MM, Koslow SH, Frazer A. Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs. *Depress Anxiety*. 1996-1997-1997;4(6):257–267.

9. Papakostas GI, Petersen T, Sklarsky KG, et al. Timing of clinical improvement and symptom resolution in the treatment of major depressive disorder. *Psychiatry Res*. 2007;149(1–3):195–200.

10. Szegedi A, Müller MJ, Anghelescu I, et al. Early improvement under mirtazapine and paroxetine predicts later stable response and remission with high sensitivity in patients with major depression. *J Clin Psychiatry*. 2003;64(4):413–420.

11. Szegedi A, Jansen WT, van Willigenburg AP, et al. 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.

12. Taylor MJ, Freemantle N, Geddes JR, et al. Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. *Arch Gen Psychiatry*. 2006;63(11):1217–1223.

13. Posternak MA, Zimmerman M. Is there a delay in the antidepressant effect? a meta-analysis. *J Clin Psychiatry*. 2005;66(2):148–158.

14. Nierenberg AA, Farabaugh AH, Alpert JE, et al. Timing of onset of antidepressant response with fluoxetine treatment. *Am J Psychiatry*. 2000;157(9):1423–1428.

15. Stassen HH, Angst J, Hell D, et al. Is there a common resilience mechanism underlying antidepressant drug response? evidence from 2848 patients. *J Clin Psychiatry*. 2007;68(8):1195–1205.

16. Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*. 1999;55(2):463–469.

17. McLachlan GJ, Peel D. *Finite Mixture Models*. New York, NY: John Wiley & Sons; 2000.

18. Muthén B, Brown H, Hunter AM, et al. General approaches to analysis of course: Applying growth mixture modeling to randomized trials of depression medication. In: Shrout PE, Keyes KM, Ornstein K, eds. *Causality and Psychopathology: Finding the Determinants of Disorders and their Cures*. New York, New York: Oxford University Press; 2011: 159–178.

19. Cinar S, Oude Voshaar RC, Janzing JG, et al. The course of depressive symptoms in unipolar depressive disorder during electroconvulsive therapy: a latent class analysis. *J Affect Disord*. 2010;124(1–2):141–147.

20. Hunter AM, Muthén BO, Cook IA, et al. Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. *J Psychiatr Res*. 2010;44(2):90–98.

21. Muthén B, Brown HC. Estimating drug effects in the presence of placebo response: causal inference using growth mixture modeling. *Stat Med*. 2009; 28(27):3363–3385.

22. Stulz N, Thase ME, Klein DN, et al. Differential effects of treatments for chronic depression: a latent growth model reanalysis. *J Consult Clin Psychol*. 2010;78(3):409–419.

23. Uher R, Muthén B, Souery D, et al. Trajectories of change in depression severity during treatment with antidepressants. *Psychol Med*. 2010;40(8): 1367–1377.

24. Wing JK, Sartorius N, Ustin TB. *Diagnosis And Clinical Measurement in Psychiatry: A Reference Manual for SCAN*. Geneva, Switzerland: World Health Organization; 1998.

25. Uher R, Maier W, Hauser J, et al. Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. *Br J Psychiatry*. 2009; 194(3):252–259.

26. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.

27. Uher R, Farmer A, Maier W, et al. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol Med*. 2008;38(2): 289–300.

28. Muthén LK, Muthen BO. *Mplus User's Guide: Statistical Analysis With Latent Variables*. 5th ed. Los Angeles, CA: Muthén & Muthén; 2008.

29. Muthén B, Asparouhov T. Growth mixture modeling: Analysis with non-Gaussian random effects. In: Fitzmaurice G, Davidian M, Verbeke G, et al, eds. *Longitudinal Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC Press; 2008:143–165.

30. Raftery AE. Bayesian model selection in social research. *Sociol Methodol*. 1995;25:111–163.

31. Schwarz G. Estimating the dimension of a model. *Ann Stat*. 1978;6(2): 461–464.

32. Nylund KL, Asparouhov T, Muthen BO. Deciding on the number of cases in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Modeling*. .2007;14: 535-569.

33. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *J Classification*. 1996;13(2):195–212.

34. Lieberman JA, Greenhouse J, Hamer RM, et al. Comparing the effects of antidepressants: consensus guidelines for evaluating quantitative reviews of antidepressant efficacy. *Neuropsychopharmacology*. 2005;30(3):445–460.

35. StataCorp. *Stata Statistical Software: Release 10*. College Station, TX: Stata Corp LP; 2007.

36. Leuchter AF, Cook IA, Gilmer WS, et al. Effectiveness of a quantitative electroencephalographic biomarker for predicting differential response or remission with escitalopram and bupropion in major depressive disorder. *Psychiatry Res*. 2009;169(2):132–138.

37. Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharm Stat*. 2008;7(2):93–106.

38. Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. *J Biopharm Stat*. 2001;11(1-2):9–21.

For Supplementary Material, go to PSYCHIATRIST.COM
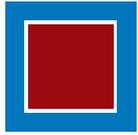
Supplementary Material

Article Title: Early and Delayed Onset of Response to Antidepressants in Individual Trajectories of Change During Treatment of Major Depression

Author(s): Rudolf Uher, PhD; Ole Mors, PhD; Marcella Rietschel, MD; Aleksandra Rajewska-Rager, MD; Ana Petrovic, PhD; Astrid Zobel, MD; Neven Henigsberg, MD; Julien Mendlewicz, MD, PhD; Katherine J. Aitchison, PhD; Anne Farmer, MD; and Peter McGuffin, PhD

List of Supplementary Material for the article

Disclaimer

This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

eFigure 1: Individual trajectories and estimated group mean. For each latent trajectory class, estimated class mean is plotted with a bold line and individual trajectories of participants belonging to each class are plotted in thin dash lines.
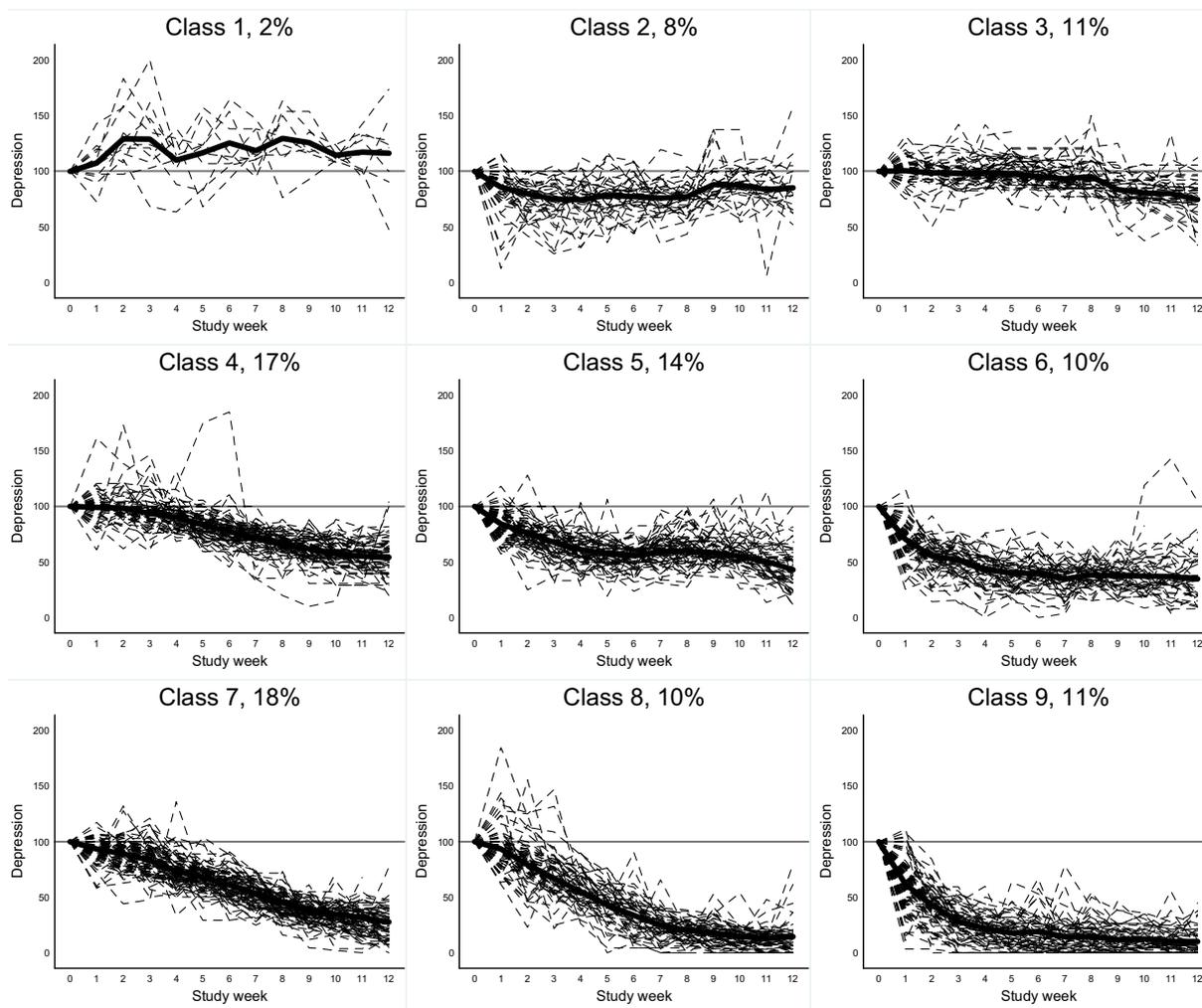
eFigure 2: Individual outliers. These three subjects had unusual course of change in depression severity during treatment and were detected in models with more than 9 classes as they formed separate classes. The first one had an early remission, but relapsed after a serious life event (sudden break-down of long term relationship) in the tenth week. The second of these subjects had a paradoxical worsening of depression with escitalopram dose increase to 20mg, the improvement continued when dose was reduced back to 10mg daily. The third outlier had a relatively unpredictable course of symptom change, but eventually experienced a remission with escitalopram dose increase to 20mg daily (and this remission persisted on 26-week follow-up, data not shown).
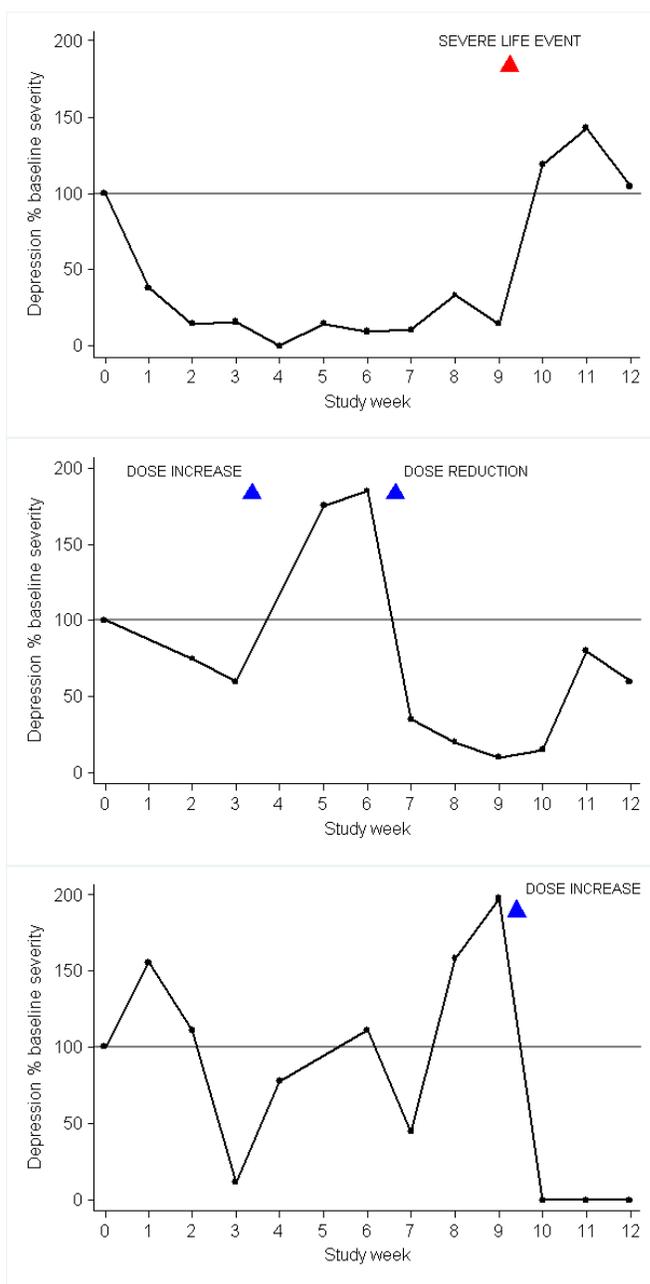
eTable 1: Longitudinal latent class model fitting. BIC = Bayesian Information Criterion, Δ BIC = decrement in BIC compared to model with one class less; BLRT = bootstrapped likelihood ratio test comparing each model with a model with one class less; Δ 2LL = 2 times the log-likelihood difference from BLRT; p – the p-value from BLRT (p value smaller than 0.05 indicates that model fits data significantly better than a model with one class less.

Classes	Parameters	Likelihood	BIC	Δ BIC	BLRT		Entropy	Proportion of individuals in each latent class									
					Δ 2LL	p		1	2	3	4	5	6	7	8	9	10
1	24	-36184	72528					1.00									
2	37	-34579	69405	3123	3210	<0.0001	0.88	0.55	0.45								
3	50	-33928	68190	1216	1303	<0.0001	0.86	0.27	0.45	0.28							
4	63	-33678	67777	413	500	<0.0001	0.84	0.16	0.38	0.31	0.16						
5	76	-33492	67493	284	371	<0.0001	0.86	0.34	0.30	0.21	0.11	0.04					
6	89	-33344	67284	209	296	<0.0001	0.83	0.02	0.16	0.28	0.20	0.19	0.14				
7	102	-33225	67133	151	238	<0.0001	0.81	0.02	0.11	0.24	0.19	0.16	0.15	0.13			
8	115	-33113	66996	137	224	<0.0001	0.82	0.02	0.10	0.23	0.16	0.15	0.13	0.11	0.11		
9	128	-33022	66900	96	183	<0.0001	0.82	0.02	0.08	0.10	0.17	0.14	0.10	0.18	0.10	0.11	
10	141	-32942	66828	73	155	<0.0001	0.83	0.02	0.08	0.10	0.17	0.14	0.10	0.18	0.10	0.10	0.002

eTable 2: Antidepressant dose titration and latent class membership. For each drug and latent trajectory class, average doses of medication at week 3, 6 and 12 are given in milligrams. Mean time to reach minimum dose and mid range dose is given in weeks.

			Escitalopram (n=464)					Nortriptyline (n=342)				
Class		n	Mean dose (mg)			Mean number of weeks to reach		n	Mean dose (mg)			Average number of weeks to reach	
			Week 3	Week 6	Week 12	Minimum dose	Mid-range dose		Week 3	Week 6	Week 12	Minimum dose	Mid-range dose
1	Worsening	73	15.71	22.86	20.00	1.14	3.43	4	58.33	66.67	.	2.00	4.00
2	Early minimal improvement	71	15.33	18.45	21.00	1.34	3.23	31	90.71	102.71	106.36	1.85	3.48
3	Minimal late improvement	56	12.07	14.41	14.63	1.26	4.41	30	73.00	91.30	90.00	2.68	5.00
4	Delayed partial response	57	12.36	14.52	16.06	1.28	4.06	62	85.85	96.59	103.03	1.63	2.93
5	Early partial response	61	13.13	16.73	20.73	1.32	4.12	54	84.46	102.56	115.58	2.04	3.54
6	Early marked response	58	12.42	14.46	15.45	1.27	3.71	44	83.54	94.74	109.82	1.79	3.31
7	Delayed marked response	62	13.33	15.63	15.33	1.24	3.51	60	86.60	92.26	98.28	1.77	2.86
8	Delayed complete remission	59	12.71	15.04	15.34	1.25	2.91	21	90.79	106.94	105.36	1.61	2.71
9	Early complete remission	53	11.41	11.81	12.50	1.40	3.79	36	82.58	92.24	92.00	1.39	3.35