

Early Symptom Trajectories as Predictors of Treatment Outcome for Citalopram Versus Placebo

Sigal Zilcha-Mano, Ph.D., Steven P. Roose, M.D., Patrick J. Brown, Ph.D., Bret R. Rutherford, M.D.

Objectives: *The high percentage of failed clinical trials for anti-depressant medications, especially in elderly populations, obscures the fact that some patients may benefit greatly from treatment. Early detection of patients who may benefit most from antidepressant medication may improve treatment decisions. We examined whether depressed patients in a large clinical trial exhibit distinct trajectories of early symptom change that predict differential response to medication or placebo.* **Methods:** *We reanalyzed data of 174 patients aged 75 years and older with unipolar depression who were randomly assigned to citalopram or placebo. We used growth mixture modeling to identify trajectories of early change (weeks 1–4) on the Hamilton Rating Scale for Depression in the citalopram and placebo conditions.* **Results:** *In the citalopram condition, two distinct trajectories of early change were detected that were associated with significantly different symptom reduction, but only one trajectory was detected for the placebo condition. One of the early trajectories of patients receiving citalopram (N = 33) showed significantly better symptomatic change than placebo; the other trajectory (N = 51) did not differ significantly from placebo. Poor baseline functional scores predicted trajectory membership, so that individuals with a score below 4.5 on baseline instrumental activities of daily living showed a higher tendency to be in the trajectory that outperformed placebo.* **Conclusions:** *The subgroup of citalopram-treated patients exhibiting better symptom trajectory early in a trial are likely to have beneficial outcomes relative to placebo. Future research should focus on developing reliable pre-treatment clinical and biological measures to identify this subgroup.* (Am J Geriatr Psychiatry 2017; 25:654–661)

Key Words: Placebo, depression, trajectories of early symptom change, personalized treatment

Major depressive disorder (MDD) in the elderly has many negative consequences, including functional decline and a higher risk for other illnesses, such as dementia, and its healthcare costs are

high.^{1–3} The current gold standard for the treatment of MDD is antidepressant medications,⁴ but even with maximal treatment many patients fail to experience sustained remission of their depression.⁵ Better

Received December 6, 2016; revised January 30, 2017; accepted February 1, 2017. From the Department of Psychology (SZ-M), University of Haifa, Israel; and the Columbia University College of Physicians and Surgeons (SPR, PJB, BRR), New York State Psychiatric Institute, New York, NY. Send correspondence and reprint requests to Dr. Sigal Zilcha-Mano, Department of Psychology, University of Haifa, Mount Carmel, Haifa 31905, Israel. e-mail: sigalzil@gmail.com

© 2017 American Association for Geriatric Psychiatry. Published by Elsevier Inc. All rights reserved.

<http://dx.doi.org/10.1016/j.jagp.2017.02.001>

pharmacological options for treating MDD are urgently needed, but a rise in failed trials of putative antidepressant agents^{6,7} has made the development of effective treatments difficult and expensive. This “psychopharmacological crisis”⁶ is especially troubling among the elderly population.

It has recently been suggested that aggregated data from many individuals may mask inter-individual variability⁸⁻¹⁰ because some patients demonstrate a clear advantage for a given medication over placebo, whereas for others this is not the case.¹¹ The complex etiology of late-life depression may result in distinct clinical depression subtypes based on their underlying biology, each requiring a different treatment approach.¹² Only some of these subtypes may respond to medication. Indeed, patient cohorts included in current antidepressant trials for the elderly show a heterogeneous response to antidepressants,¹³ which may have contributed to treatment failure for the cohort as a whole. Such failures increase the cost of drug development, delay marketing, and eventually limit treatment for patients who could benefit from the medication. Early detection of responsive patients can aid in decision-making, improve response rates, and lower costs by focusing on patients who could benefit most from treatment and referring others to alternative treatment options.

Advanced statistical tools can help identify distinct trajectories of change in different subpopulations within the same cohort. In a re-analysis of the data from a clinical trial of duloxetine, Gueorguieva et al.¹⁴ identified distinct trajectories for responders (76.3% of the sample) and nonresponders (23.7% of the sample) in an antidepressant-treated subsample, whereas placebo-treated patients were characterized by a single response trajectory. Patients in the “responders” trajectory had better treatment outcomes than the placebo group, whereas those in the “non-responders” trajectory had poorer outcomes than the placebo group. Although these findings are promising, they are currently restricted to younger populations, and have not yet been studied in the elderly. Furthermore, trajectories of change were evaluated based on data that is available only at the end of treatment, limiting the clinical usefulness of identifying non-responders. Evaluating early trajectories of change as predictors of treatment outcome is more relevant for practical use.

In the present study we conducted a secondary analysis of a randomized controlled trial (RCT)

comparing medication with placebo in patients diagnosed with unipolar depression, aged 75 years and older. This large, well-conducted study failed to find significant outcome differences across the entire sample between participants randomized to citalopram and those receiving placebo. In the current study, however, we were interested in determining whether it is possible to identify a significant medication-responsive subgroup of participants based on distinctive trajectories of early symptomatic change. We chose to focus on the first four symptom assessments, because most previous reports have used between three¹⁵ and five¹⁶ sessions for detecting early change in treatment. We reasoned that choosing the first four sessions can produce clinically meaningful information early enough to affect subsequent treatment, including the possibility of switching to a different treatment when necessary.¹⁷

METHODS

Sample and Clinical Trial Procedures

The procedures used in this multi-site, placebo-controlled RCT have been previously described.¹³ Briefly, 174 community-dwelling men and women aged 75 years or older who met DSM-IV criteria (based on a Structured Clinical Interview for DSM-IV Axis I Disorders interview) for non-psychotic unipolar depression (single or recurrent), with a baseline 24-item Hamilton Rating Scale for Depression (HRSD¹⁸) score of 20 or higher, participated in this 8-week RCT. All patients began the trial with a one-week, single-blind placebo lead-in, with the baseline visit conducted at the end of the lead-in period. Patients were randomized in 15 centers to citalopram (20 mg/day) or matched placebo at a ratio of 1:1 only if they continued to meet inclusion and exclusion criteria at the end of the placebo lead-in period. At the end of the fourth week, patients with an HRSD score less than 10 had their medication dose increased to two pills per day (i.e., 40 mg of citalopram, or two placebo pills). Clinical assessments were conducted at baseline and at weeks 1, 2, 3, 4, 5, 6, and 8 (final week). For this analysis, baseline and weekly assessments of the HRSD were used, together with the intake assessment of the instrumental activities of daily living (IADL¹⁹).

Statistical Analyses Overview

Identifying Early Trajectories of HRSD Scores Across Treatment

To identify distinct trajectories of HRSD scores across early treatment (weeks 1–4), we used growth mixture modeling^{20–22} with the *lcm* package of the R software environment.^{23,24} To examine whether distinct early trajectories emerged for patients in the citalopram and the placebo conditions, we tested the fit of separate models for each condition. We considered linear, quadratic, square root, and cubic trends over time, with one to three trajectory classes. The selection of the best model was based on the Akaike information criterion (AIC) and the Lo-Mendell-Rubin (LMR) likelihood ratio test.²⁵ LMR is used to select the number of classes by evaluating whether a model with one class fewer than the fitted model describes the data as accurately as the fitted model. We assessed classification accuracy using the entropy value, ranging between 0 and 1, with values closer to 1 corresponding to better classification accuracy.

Note that the LMR likelihood ratio test (LRT) statistic does not have a χ^2 distribution because of boundary conditions, but its distribution can be determined empirically by bootstrapping. First, we estimated both the k-class and the k-1-class model and calculated the observed LMR LRT value. Next, we generated 100 samples using the parameter estimates from the k-1-class model, and for each generated sample we obtained the log likelihood value for both the k-1 and the k-class model, enabling us to compute the LRT values for all generated samples. Finally, we calculated the p value for the LRT by comparing its observed value to the estimated null distribution.

Prediction of Early Trajectory Based on Pre-Treatment Characteristics

We conducted decision tree analyses²⁶ with the R *party* package,²⁷ using random forest variable selection²⁸ and Monte Carlo simulation for multiple-testing adjustment.²⁹ We used the following baseline characteristics as potential predictors of trajectory assignment in the citalopram condition: age, sex, education, age of first MDD episode, symptom severity (assessed as the Hamilton score at intake), baseline instrumental activities of daily living (assessed using the IADL), and study site.

Trajectory Assignment as a Predictor of HRSD Reduction Across Treatment and Clinical Response

Because the separate analysis revealed two early trajectories of the subsample receiving citalopram, and only one trajectory of the subsample receiving placebo, mixed-model analyses³⁰ were performed to assess whether patients receiving placebo had significantly different responses from those receiving medication, who were classified into responder or non-responder trajectories. The outcome variable in the mixed model was HRSD across treatment. The predictors were trajectory class membership (placebo, medication trajectory responder, or medication trajectory non-responder), time, and their interaction. To control for the potential confounding of baseline covariates when comparing the randomized placebo group with the nonrandomized trajectory responder and non-responder groups, we used a propensity scoring approach. We calculated the predicted probability (propensity score) for each patient to be in the trajectory non-responder class, and used this probability as a covariate in the mixed model. We controlled for site because of the differences that were found between sites in the main outcome report.¹³ Trajectory responders were defined as patients who were classified in the responder trajectory at week 4. Clinical responders were defined as patients with at least 50% improvement over baseline HRSD scores at the end of treatment (week 8). We used the χ^2 test to test independence and the odds ratio (OR) for assessing the relationship between trajectory response and clinical response.

RESULTS

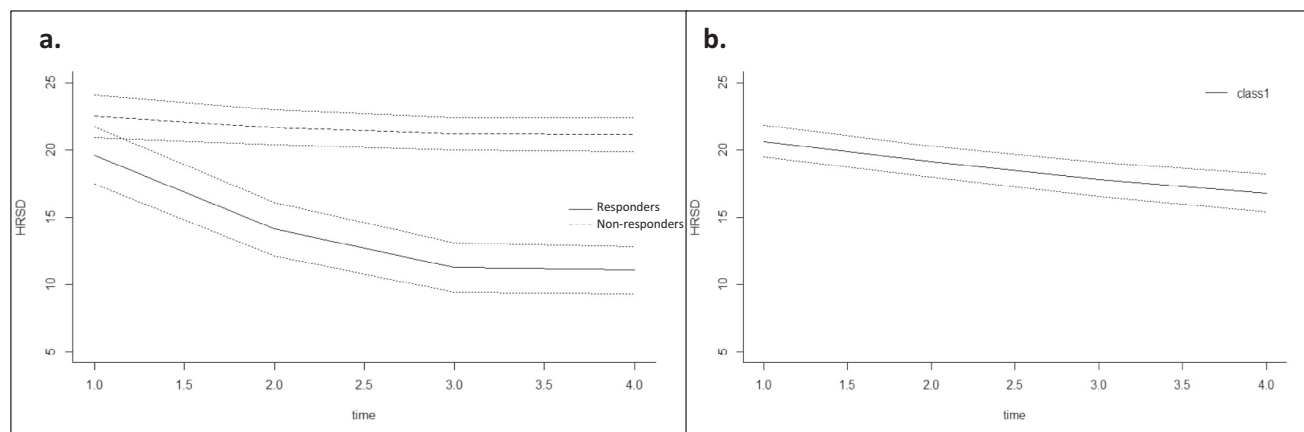
Clinical and Demographic Characteristics of Participants

Eighty-four patients were randomized to citalopram and 90 to placebo, of whom 58% were women; mean age was 79.6 years (SD: 4.4 years), and the mean baseline HDRS score was 24.3 (SD: 4.1).

Identifying Trajectories of HRSD Across Treatment

According to the AIC and the LMR, the best-fitting set of models for the citalopram and placebo arms were quadratic growth mixture models (see [Tables S1 and](#)

FIGURE 1. Estimated means of early trajectories for [a] patients receiving medication and [b] patients receiving placebo, with 95% confidence interval.



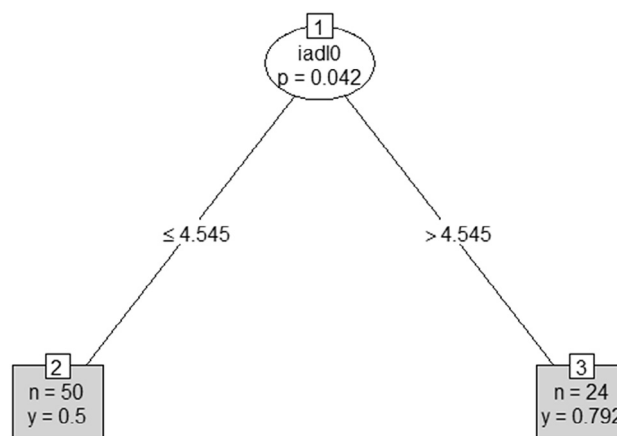
S2 in the online Supplemental Material). For the citalopram arm, the model with two classes fit the data best. Figure 1a presents the estimated means for the two trajectory classes in the citalopram arm. The first class, with 39.3% probability of membership, can be interpreted as the class of trajectory responders. The second class, with 60.7% probability of membership, can be interpreted as the class of trajectory non-responders. Based on the AIC, there was no evidence in the present data of more than one trajectory class in the placebo arm (Figure 1b).

Prediction of Trajectory Membership

The tree decision analysis revealed a significant split in the IADL variable ($t_{(74)} = 2.63$, $p = 0.04$), such that patients with a score below or equal to 4.54 on baseline IADL showed a higher tendency to be in the trajectory that outperformed placebo. No other pre-treatment characteristics were significant predictors of trajectory membership. Figure 2 presents the tree for trajectory assignment.

We conducted another analysis, in which in addition to all other pre-treatment variables we also added the following variables: (a) the Cumulative Illness Rating scale for Geriatrics; (b) the Mini-Mental State Examination; and (c) the Stroop Color and Word test, a response inhibition task assessing a central component of the executive functions (Sneed et al.³¹). Findings were similar to the original analysis, and none of these three

FIGURE 2. Decision-tree learning for trajectory assignment, based on the random forest algorithm for variable selection, and Monte Carlo simulation for multiple testing adjustment. Y: proportion of non-responders.



variables were chosen as robust predictors. Future studies should further investigate these issues.

Trajectory Membership as a Predictor of Change in HRSD Across Treatment and Clinical Response

The linear mixed models comparing HRSD scores over time for trajectory responders receiving active drug,

trajectory non-responders receiving active drug, and patients receiving placebo showed a significant interaction between trajectory membership and time ($F_{(2,1125)} = 7.24$, $p = 0.0008$). The two citalopram trajectories showed significantly different outcome reduction ($B = -0.42$, $SE = 0.13$, $t_{(1125)} = -3.12$, $p = 0.001$). Trajectory responders receiving active drug showed greater reduction in HRSD scores than did patients receiving placebo ($B = -0.48$, $SE = 0.13$, $t_{(1125)} = -3.66$, $p = 0.0003$). No significant differences were found between trajectory non-responders receiving active drug and patients receiving placebo ($B = -0.05$, $SE = 0.11$, $t_{(1125)} = -0.47$, $p = 0.63$).

For the medication trajectories, using the definition of clinical response as a 50% or greater improvement from baseline HRSD at the completion of treatment, trajectory responders had a greater chance of being clinical responders ($\chi^2_{(1)} = 28.08$, $OR: 14.5$, $p < 0.0001$). The correspondence between trajectory responders and clinical responders indicated that 25 of 33 (75.75%) trajectory responders were also clinical responders, and 42 of 51 (82.35%) trajectory non-responders were classified as clinical non-responders.

Sensitivity Analysis

We assessed the effect of missing data on our results by performing a sensitivity analysis under missing not at random (MNAR) assumptions. The linear mixed model assumes that the missing data are missing at random (MAR).³² When this assumption does not hold, the longitudinal process and the missing data process can be simultaneously modeled in a so-called joint model.²⁴ There is a close association between drop-out from therapy and joint modelling in that patients' drop-out time can be considered as a survival outcome.³³ Frequently, a longitudinal process is associated with a survival process (in our case, patient dropout), and the joint model captures this correlation to provide a valid inference. We applied a joint model to our data, performing the analysis based on an MNAR assumption as a sensitivity analysis. The model revealed the same pattern of findings reported by assuming MAR, supporting the robustness of our findings vis-à-vis possible missing data effects.

DISCUSSION

The present study sought to determine whether distinct early trajectories of change in a cohort of elderly patients receiving antidepressant medication can predict their ability to benefit from antidepressant medication compared with placebo. Findings suggest that although no significant differences were found in the primary study that compared the citalopram and the placebo conditions for the cohort as a whole, two distinct early trajectories of change were detected in patients receiving citalopram—one with a significantly better outcome than placebo, and one that did not differ significantly from placebo. For patients receiving citalopram, baseline IADL could predict assignment to one of the two distinct early trajectories of change, so that patients with a score of 4.5 or less on the baseline IADL showed a higher tendency to be in the trajectory that outperformed placebo.

These findings are consistent with a previous study that focused on a younger population and also identified two distinct trajectories of change in patients receiving antidepressant medication, only one of which outperformed placebo.¹⁴ The presence of two distinct trajectories in the medication group suggests that the patient cohort included in the current trial was heterogeneous in its response to treatment, and that using the average value for this variable may have led to treatment failure for the cohort as a whole. Moreover, our focus on early trajectories of change rather than simply the quantity of early change (e.g., $\geq 25\%$ improvement after four weeks³⁴) may more accurately identify individuals most likely to benefit from antidepressant treatment. Detecting patients likely to respond to citalopram early during treatment may enable clinicians to continue treatment for those who will benefit from it, and to refer others to alternative treatments.

In the analysis of Gueorguieva et al.,¹⁴ which focused on a younger population, 76.3% of the sample were classified in the responder trajectory that outperformed placebo, whereas in the present study, which focused on the elderly population, only 39.3% were classified in the responder trajectory that outperformed placebo. Although these percentages are not entirely comparable, the finding may suggest potential differences between the elderly and younger populations. This rationale may explain the higher percentages of failed trials in the former group, and it is consistent

with the view of depression in the elderly as a heterogeneous neuropsychiatric syndrome in later life, caused by pathophysiologically distinct brain disorders and resulting in common behavioral manifestations.³⁵

Identification of different treatment trajectories raises the possibility of using patients' pre-treatment characteristics to predict their future course. In the present study, we found that a baseline IADL score less than 4.5 was associated with the trajectory that outperformed placebo. Whereas in younger populations baseline HRSD levels play a crucial role in determining who may benefit most from antidepressants,³⁶ in older populations it is the IADL that plays such a role. The importance of functional status is consistent with phenomenological differences observed in late life depression as compared to MDD in younger patients. MDD in the elderly is characterized less by depressed mood³⁷ and more by functional complaints, such as sleep disturbance, fatigue, and psychomotor retardation,³⁸ as well as by poor memory and concentration, slow cognitive processing speed, and executive dysfunction,³⁹ when compared with younger adults with the same condition. Future studies should further explore the role of IADL levels in determining who may benefit most from antidepressants, and explore other relevant characteristics of those with lower levels of IADL.

Another interesting finding of this study was that only one trajectory could be found for the patients randomized to placebo. It is possible that individuals who responded during the one-week lead-in placebo washout period, and were therefore excluded from the trial before randomization, could have formed a distinct trajectory, although they were too few in number to create a meaningful subgroup (N = 6). The fact that patients in the placebo condition were all assigned to one trajectory may suggest that although variations exist between patients receiving placebo, overall they are best perceived as one group of patients, showing similar patterns of change. Another possibility is that the sample size was not large enough to enable the detection of subgroups of patients with more subtle differences within the placebo condition, although the present findings are consistent with Gueorguieva et al.'s study,¹⁴ which had a much larger sample size, demonstrating a single trajectory for the placebo condition.

The current findings are limited by the moderate sample size. Validation of the present findings with

larger sample sizes and follow-up data sets are needed. A crucial path for future research is to evaluate whether these findings are specific to citalopram or to antidepressant medications in general. If it is indeed found that the same patients are in the non-responder groups for different drugs, these patients should be offered psychotherapy or other nonpharmacological treatment. If, however, patients who are non-responders for one medication are found to respond positively to another, this information can have important implications for personalized medicine. To test such a possibility, a more complex study design is needed. If non-responder trajectories are medication-specific, it is necessary to examine whether an earlier initiation of a second medication would result in greater benefit and lower cost than continuing with the first one.

Our study is also limited by the use of a single dose of citalopram. It is possible to speculate that trajectory shape is different in dose-escalation designs in which study participants do not receive a therapeutic dose of medication until partway through the study. More specifically, responder and non-responder trajectories may diverge later, so that clinicians wishing to utilize these results in practice may need to delay the assessment of whether patients have had an early response (i.e., to determine trajectory membership and thus whether to switch or augment treatment strategy). Nevertheless, Gueorguieva et al.,¹⁴ in their analysis of duloxetine studies that did use dose escalation designs, reported findings consistent with our results.

An important implication of the present findings is that there was a subgroup of citalopram-treated patients who did experience significant improvement relative to placebo, although the overall parent RCT was negative and citalopram did not differentiate from placebo. This suggests that citalopram may be a useful treatment for some patients, but that characteristics of these patients need to be identified before their enrollment in future confirmatory RCTs. In this way, the present findings are consistent with other analyses demonstrating that comparing overall mean differences between drug and placebo groups does not account for different trajectories of treatment response¹⁴ and is relatively insensitive to large and significant changes experienced by subgroups of the sample.¹¹ Second, the present findings add to other published reports suggesting that patients who do not demonstrate at least a 30% improvement in baseline symptomatology by 4 weeks are unlikely to respond

to medication by 8 weeks, and may profitably be switched to another treatment at the 4-week time point.⁴⁰ If the present findings can be generalized to clinical practice, psychiatrists may wish to consider switching or augmentation of treatment for individuals not showing a desirable trajectory of change.

This is the first study to examine systematically the effect of early trajectories of symptomatic change as predictors of outcome in elderly patients with MDD. In a negative trial, we detected two distinct early trajectories of change for elderly patients receiving citalopram for major depressive disorder, only one of which was associated with a significantly better outcome than placebo. It may be the case that analyzing mean depressive symptom change across all patients assigned to medication versus placebo may obscure large and meaningful drug-placebo differences in patient subgroups. These findings thus

demonstrate the promise of investigating distinct early trajectories for signal detection in antidepressant clinical trials. Clinically, characterizing early patterns of change may contain useful information about whether to continue an antidepressant trial and thus diminish the length of time depressed patients are exposed to ineffective treatments.

Work on this paper was supported by R01 MH102293 (BRR) and T32 MH015144 (SPR). Dr. Rutherford reports receiving consulting fees from Pfizer. Drs. Zilcha-Mano, Brown, and Roose have no disclosures to report.

APPENDIX: SUPPLEMENTARY MATERIAL

Supplementary data to this article can be found online at [doi:10.1016/j.jagp.2017.02.001](https://doi.org/10.1016/j.jagp.2017.02.001).

References

1. Barnes DE, Yaffe K, Byers AL, et al: Midlife vs late-life depressive symptoms and risk of dementia: differential effects for Alzheimer disease and vascular dementia. *Arch Gen Psychiatry* 2012; 69:493-498
2. Frasure-Smith N, Lesperance F, Talajic M: Depression and 18-month prognosis after myocardial infarction. *Circulation* 1995; 91:999-1005
3. Morichi V, Dell'Aquila G, Trotta F, et al: Diagnosing and treating depression in older and oldest old. *Curr Pharm Des* 2015; 21:1690-1698
4. Hollon SD, Thase ME, Markowitz JC: Treatment and prevention of depression. *Psychol Sci Public Interest* 2002; 3:39-77
5. Rush AJ, Trivedi MH, Wisniewski SR, et al: Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 2006; 163:1905-1917
6. Cressey D: Psychopharmacology in crisis. *Nat News* 2011; doi:10.1038/news.2011.367. 201110
7. Khan A, Brown WA: Antidepressants versus placebo in major depression: an overview. *World Psychiatry* 2015; 14:294-300
8. DeRubeis RJ, Cohen ZD, Forand NR, et al: The personalized advantage index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS ONE* 2014; 9:e83875
9. Fisher AJ: Toward a dynamic model of psychological assessment: implications for personalized care. *J Consult Clin Psychol* 2015; 83:825-836
10. Muthén B, Brown HC: Estimating drug effects in the presence of placebo response: causal inference using growth mixture modeling. *Stat Med* 2009; 28:3363-3385
11. Thase ME, Larsen KG, Kennedy SH: Assessing the "true" effect of active antidepressant therapy v. placebo in major depressive disorder: use of a mixture model. *Br J Psychiatry* 2011; 199:501-507
12. Valengo LCL, Florindo S, Orestes VF: Mood disorders in the elderly: prevalence, functional impact, and management challenges. *Neuropsychiatr Dis Treat* 2015; 12:2105-2114
13. Roose SP, Sackeim HA, Krishnan KRR, et al: Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. *Am J Psychiatry* 2004; 161:2050-2059
14. Gueorguieva R, Mallinckrodt C, Krystal JH: Trajectories of depression severity in clinical trials of duloxetine: insights into antidepressant and placebo responses. *Arch Gen Psychiatry* 2011; 68:1227-1237
15. Haas E, Hill RD, Lambert MJ, et al: Do early responders to psychotherapy maintain treatment gains? *J Clin Psychol* 2002; 58:1157-1172
16. Lutz W, Hofmann SG, Rubel J, et al: Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *J Consult Clin Psychol* 2014; 82:287-297
17. Grilo CM, White MA, Masheb RM, et al: Predicting meaningful outcomes to medication and self-help treatments for binge-eating disorder in primary care: the significance of early rapid response. *J Consult Clin Psychol* 2015; 83:387
18. Hamilton M: A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; 23:56-62
19. Lawton MP: The functional assessment of elderly people. *J Am Geriatr Soc* 1971; 19:465-481
20. Verbeke G, Lesaffre E: A linear mixed-effects model with heterogeneity in the random-effects population. *J Am Stat Assoc* 1996; 91:217-221
21. Muthén B, Shedden K: Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 1999; 55: 463-469
22. Proust C, Jacqmin-Gadda H: Estimation of linear mixed models with a mixture of distribution for the random effects. *Comput Methods Programs Biomed* 2005; 78:165-173
23. Proust-Lima C, Philipps V, Diakite A, et al: Lcmm-package. Available at <https://cran.r-project.org/web/packages/lcmm/lcmm.pdf>. Accessed 2015.
24. Proust-Lima C, Philipps V, Liqueur B: Estimation of extended mixed models using latent classes and latent processes: The R package lcmm. *arXiv preprint arXiv:1503.00890*. Accessed 2015. 2015

25. Lo Y, Mendell NR, Rubin DB: Testing the number of components in a normal mixture. *Biometrika* 2001; 88:767-778
26. Hothorn T, Hornik K, Van De Wiel MA, et al: A lego system for conditional inference. *Am Statistician* 2006; 60:257
27. Hothorn T, Hornik K, Zeileis A: Unbiased recursive partitioning: a conditional inference framework. *J Comput Graph Stat* 2006; 15:651-674
28. Strobl C, Malley J, Tutz G: An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods* 2009; 14:323
29. Strasser H, Weber C: On the asymptotic theory of permutation statistics. Available at <http://epub.wu.ac.at/102/>. 1999
30. Raudenbush SW, Bryk AS: *Hierarchical Linear Models: Applications and Data Analysis Methods*. Vol. 1. Thousand Oaks, CA: Sage, 2002
31. Sneed JR, Culang ME, Keilp JG, et al: Antidepressant medication and executive dysfunction: a deleterious interaction in late-life depression. *Am J Geriatr Psychiatry* 2010; 18:128-135
32. Little RJ: Modeling the drop-out mechanism in repeated-measures studies. *J Am Stat Assoc* 1995; 90:1112-1121
33. Asar O, Ritchie J, Kalra PA, et al: Joint modelling of repeated measurement and time-to-event data: an introductory tutorial. *Int J Epidemiol* 2015; 44:334-344
34. Hirschfeld RM, Montgomery SA, Aguglia E, et al: Partial response and nonresponse to antidepressant therapy: current approaches and treatment options. *J Clin Psychiatry* 2002; 63:826-837
35. Rutherford RR, Taylor WD, Brown PJ, et al: Biological aging and the future of geriatric psychiatry. *J Gerontol A Biol Sci Med Sci* 2016; doi:10.1093/gerona/glw241
36. Fournier JC, DeRubeis RJ, Hollon SD, et al: Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010; 303:47-53
37. Krishnan KRR, Hays JC, Tupler LA, et al: Clinical and phenomenological comparisons of late-onset and early-onset depression. *Am J Psychiatry* 1995; 152:785-788
38. Christensen H, Jorm A, Mackinnon A, et al: Age differences in depression and anxiety symptoms: a structural equation modelling analysis of data from a general population sample. *Psychol Med* 1999; 29:325-339
39. Butters MA, Whyte EM, Nebes RD, et al: The nature and determinants of neuropsychological functioning in late-life depression. *Arch Gen Psychiatry* 2004; 61:587-595
40. Sackeim HA, Roose SP, Lavori PW: Determining the duration of antidepressant treatment: application of signal detection methodology and the need for duration adaptive designs (DAD). *Biol Psychiatry* 2006; 59:483-492