Published in final edited form as:

Am J Geriatr Pharmacother. 2011 December ; 9(6): 483–494. doi:10.1016/j.amjopharm.2011.09.001.

Trajectory Classes of Depression in a Randomized Depression Trial of Heart Failure Patients: a Reanalysis of the SADHART-CHF Trial

Maragatha N. Kuchibhatla, PhD^{1,2} and Gerda G. Fillenbaum, PhD^{1,3}

¹Center for the Study of Aging and Human Development, Duke University Medical Center, Durham, NC

²Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC

³Geriatrics Research, Education, and Clinical Center, Veterans Administration Medical Center, Durham NC

Abstract

Objective—To determine whether, in drug intervention trials, growth mixture modeling (GMM) is able to identify drug responsive trajectory classes that are not evident under traditional growth modeling approaches.

Study Design and Setting—We re-analyzed acute phase (bi-weekly data on up to seven occasions) and longitudinal (12 months) data on the 469 patients in the SADHART-CHF study of the safety and efficacy of sertraline for depression in patients with heart failure. GMM was used to identify the trajectory classes present in the treatment and placebo groups, based on Hamilton Depression Rating Scale (HDRS) scores.

Results—Two distinct trajectory classes were identified in the treatment group – (1) chronic depressives (12%), who remained depressed through the treatment phase, (2) responders (88%), who had scores indicating nondepression at acute phase conclusion. At baseline, chronic depressives were distinguished by higher HDRS score, presence of implantable cardioverter defibrillators, and a history of anxiety. During followup, they were more likely to have unstable angina. Only responders remitted (70%). Three distinct trajectories were identified in the placebo group – (1) moderating depressives (19%), (2) temporary improvers (13%), (3) responders (68%). At baseline, the classes differed in mean HDRS score, responders' scores falling between the other two classes, and proportion with renal disease. Only remission differed at followup -- responders (76%), moderating depressives (21%), temporary improvers (3%).

Where the traditional analytic approach found improvement from moderate to mild depression, but no significant treatment effect, GMM found response in 20% more people in the treatment compared to the placebo group.

Conclusions—Unlike conventionally used standard analytic approaches which focus on intervention impact at study end or change from baseline to study end, GMM enables maximum

© 2011 Excerpta Medica, Inc. All rights reserved.

Corresponding Author: Maragatha Kuchibhatla, Box 3003, DUMC, Durham, NC 27710. mnk@geri.duke.edu.

The clinical trials registration number in 'clinicaltrials.gov' is NCT00078286.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

use of repeated data to identify unique trajectories of latent classes which are responsive to the intervention.

Keywords

latent class trajectories; depression; drug trials

Introduction

Among patients with cardiovascular disease, estimates of the presence of comorbid depression range from 15%–20%. Depression is a risk factor for additional morbidity and functional impairment, as well as for mortality, particularly if onset of major depression follows cardiovascular disease [1–7]. In order to try to prevent such adverse outcomes, antidepressants may be prescribed when depression is recognized. However, as review of major studies indicates, antidepressants are not always effective [8], although when the initial antidepressant is ineffective, an alternative may work better [9,10].

Thus, different patients respond to different antidepressants. What we do not know, but would like to know, preferably before prescribing, are the patient characteristics that determine responsivity to a particular antidepressant. A recent excellent review [11] indicates the extensive difficulties in this area. Clinical characteristics, psychiatric comorbidities, potential genetic markers, or measures of brain functioning and metabolism were rarely useful in consistently identifying those antidepressants uniquely effective for patients with major depressive disorder. Our specific focus is congestive heart failure (CHF) patients with major depression, since they are at particular risk of adverse outcomes. A recent study of CHF patients was able to show only that those who were more depressed at baseline, or who had additional hospitalizations for CHF after entry into the study, were less likely to remit [12]. While concentration has been on determining whether an antidepressant is efficacious, less work has been done on identifying subgroups for which an antidepressant may be particularly effective, although statistical approaches to do this are available. For example, tree-based models, designed for cross-sectional data, can identify subgroups, based on a combination of variables that are effective predictors of outcomes; in regression models interaction terms that identify subgroups can be tested.

The traditional growth modeling approach (random-effects/multilevel modeling), commonly used for FDA approval, assumes that a single group of growth parameters fits the population and that the covariates that affect the growth factors affect each individual in a similar way. While a subgroup by growth parameter interaction can be included in such models if the subgroup is known *a priori*, it cannot be included in these models if the subgroup is unknown and exploratory. Fitting a single linear representation of change over time may obscure different patterns that better describe distinct subgroups. To address this concern, we use growth mixture modeling [13, 14]. We re-analyze data obtained in the SADHART–CHF (Sertraline Against Depression and Heart Disease in Congestive Heart Failure) study [15,16], since in this study, outcome using standard analytic techniques has been published and the data are available. Using the standard statistical approach, SADHART-CHF found that patients on the antidepressant sertraline improved from a moderate to a mild level of depression, but this improvement differed little from that of patients in the placebo group. The current re-analysis, using growth mixture modeling [13,14], is intended to ascertain whether unique subgroups were present in the treatment and placebo groups; whether, in the treatment group, the characteristics of the subgroups differed from each other in a manner permitting identification of persons responsive to sertraline; and whether, within the treatment and placebo groups, the latent class trajectories identified differed in remission

status (a preset level of improvement in measured depression) and in predictors of clinical outcomes.

Methods

Sample

The SADHART-CHF study is an NIMH-sponsored, prospective, randomized, double-blind, placebo-controlled trial, designed to assess the safety and efficacy of sertraline in the treatment of heart failure (HF) patients with major depressive disorder (MDD). To enhance the generalizability of the findings, study participants were recruited from an academic center, and two community hospitals and clinics. Data were gathered between August 2003 and March 2008. The clinical trials registration number in 'clinicaltrials.gov' is NCT00078286.

The study design [15] and the results of the primary outcomes [16] have been published. Briefly, the study population consisted of patients who were at least 45 years old with a diagnosis of chronic HF of any etiology. Eligible patients had to meet the Diagnostic and Statistical Manual of Mental Disorders IV [17] criteria for MDD and have a history of chronic systolic HF characterized by left ventricular ejection fraction < 45% and New York Heart Association (NYHA) class II. Qualified patients (n=469) were randomized to one of two treatments groups: sertraline or placebo.

The trial included 2 phases: (1) a double blind treatment phase (acute phase) -- data were gathered at baseline, and then as close as possible to weeks 2, 4, 6, 8, 10, and 12, but stretching to 14 weeks to accommodate patients unable to maintain the biweekly schedule; and (2) a long-term open treatment follow-up phase. Patients started with an initial dose of sertraline or a placebo pill of 50mg/day, titrated up bi-weekly by 50 mg/day to a maximum of 200 mg/day. If a patient was unable to tolerate the increment dosages, increase in dosage was delayed over the rest of the acute phase.

At baseline, information was collected on demographics, hospitalizations in the past year, history of cardiac procedures, medical conditions (renal, diabetes), cardiovascular and psychiatric history, and cardiac medications. Information was updated biweekly, including performance on the 17-item Hamilton Depression Rating Scale (HDRS) [18] and all medications. The HDRS is a semi-structured interview administered by a trained interviewer, that assesses different aspects of affective disorders (e.g., type of depression, insomnia, anxiety, somatic symptoms). Depending on the item, scoring may be on a 3-point (absent to clearly present), or 5-point (absent to severe) scale. The potential scoring range is 0–54, the cut-point indicative of depression used here was > 7, with remission defined as either attainment of a score < 8 at study end or decline to half the HDRS baseline score. Following the acute phase, all subjects were contacted at 6 months, one year, and annually thereafter, and information obtained on any cardiac or psychiatric events, hospitalizations and emergency room visits in the interim. All serious adverse events were reported as and when they occurred. All intervening events were adjudicated by cardiologists not associated with the study.

The primary end points for the acute phase were (1) change in HDRS score, and (2) cardiovascular status categorized as improved, unchanged, or worsened. Focus in the present paper is on subgroups of change in HDRS score and possible determinants of such change.

Statistical Analysis

The original trial used an intention-to-treat analysis [15,16], i.e., all participants enrolled in the study were included. A random coefficients regression model was used to look at the trajectory of change in HDRS scores, and to compare treatment and placebo groups. This procedure has the advantage of allowing information to be used to the extent that it is available, and reduces concern over missing data. The model included the fixed effects of treatment, the natural log of time and time squared and their interactions with treatment, as well as the random effects of patient, interaction of time by patient and by patient squared, and site [16]. Overall differences in response to treatment were tested using chi square (and the Mantel-Haenszel test when clinical site was controlled). Thus, focus was on ascertaining whether outcome in the treatment and placebo groups differed significantly in response to the intervention (sertraline). Focus was not on determining whether a particular class of patients was responsive to sertraline.

Unlike traditional growth modeling techniques which assume that only a single class is present, GMM relaxes this single population assumption, and tests if more than one distinct class can be used to describe the data. That is, it checks whether there are underlying, meaningful, subgroups of trajectories (latent classes, identified by different growth curves) in the dataset. So, instead of considering individual variation around a single mean growth curve, GMM allows different classes of individuals to vary around different mean growth curves. Further, GMM provides a means for identifying characteristics that distinguish among these subgroups [13,14]. While we have implemented GMM using Mplus [19], it is also possible to use Proc Traj (in SAS 9.2), which operationalizes an approach described by Nagin [20]. In that approach, which differs slightly from the one used here, the growth trajectories within each class are assumed to be homogeneous, i.e., the parameters in each class are fixed. The GMM approach, on the other hand, assumes different growth parameters across classes [21,22]. As with the random coefficients model, GMM allows data to be used to the extent that it is available (i.e., the model is fit under the assumption of missing at random), and reduces concern over missing data.

The proposed mixture model is shown in Figure 1. The components of this model are the latent class variable *c*, repeated continuous outcomes of HDRS, and the growth components (intercept [I], linear [S], and quadratic [Q]) of the model. This model can be extended to include covariates influencing the growth components, and latent classes influencing distal outcomes.

The trajectory classes of HDRS are identified separately for the treatment and the placebo group, and within each group are compared with respect to demographic and other baseline characteristics, as well as acute phase and one year survival, cardiac related events, and drop out (details given in Tables 1–4). The aim of these analyses is to determine the characteristics of the various trajectory classes of HDRS, and to examine the characteristics of the class that has benefited the most from the trial. The variables that were selected for comparisons between the classes were based on the characteristics used in the SADHART-CHF primary analysis paper [16], defined according to the study protocol. Differences in baseline and post-treatment variables by trajectory classes were compared using non-parametric Wilcoxon-rank sum test or the Kruskal-Wallis test.

Fit statistics used to determine the number of classes for the model include Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), subject-specific Bayesian Information Criteria (SSBIC), Lo-Mendell-Rubin test, and bootstrap likelihood ratio test. In comparison of alternative or successive models, lower values of AIC, BIC and SSBIC are indicative of better fit [23,24]. In addition, entropy and conditions numbers (which range

from 0–1, higher values imply better fit) were used to determine the number of classes that best fit the data. To determine the form of the model, we fit linear, quadratic and piece-wise linear models to the treatment group, placebo group, and the combined sample. Piecewise linear had the best fit. Descriptive statistics used to characterize the sample were carried out using SAS, version 9.2 (SAS Institute, Cary, NC). Trajectory class analysis was carried out using M-plus software version 5.2 [19]. This study was conducted with the consent of the Duke University Medical Center Institutional Review Board.

Results

The characteristics of the entire sample have been presented previously [16]. Briefly, average age was 62.1 years, 40.5% were female, 56.9% were White, 28.1% were classed as NYHA Class II, 47.5% Class III, and 24.3% Class IV. Findings from the original study indicated comparable scores at baseline for the treatment and placebo groups (mean total HDRS score 18.3 for each, with standard deviation (sd) 5.5 in the treatment and 5.4 in the placebo group). The adjusted random coefficients regression model showed a significant decline in total HDRS score for each group, but the difference in decline between the two groups, was not significant.

GMM analysis findings for the treatment and placebo groups, separately by trajectory class, are based on 467 patients, rather than the original 469, since HDRS was missing on all visits for two patients.

Treatment group

A two class piece-wise linear model best fit the treatment group (Figure 2). The first class, a chronic depressive class (n=28; 12% of the treatment sample), had an estimated mean HDRS score at baseline (the intercept) of 20.9. This trajectory had a post-randomization drop at week 2 to an HDRS score of about 16.9, with HDRS scores stable from week 2 and ending at week 14 with an HDRS depression score of about 17.6. A second class, a responder class (n=205; 88% of the treatment sample) had a post-randomization drop at week 2 from a baseline (intercept) HDRS score of 17.8 to about 8.6, with continued stable HDRS depression scores from week 2 to week 14 when the lowest HDRS score of 5.5 was reached.

Comparison of the baseline demographic, hospitalization, cardiovascular and psychiatric history of the two treatment group trajectory classes (Table 1) shows that the mean ages of the chronic depressive class and the responder class were 61.9 and 63.1 years respectively, but were not significantly different (p-value = 0.3940). There were no significant differences in race (Whites: 71.4% vs nonwhite (all but 2 African American) 54.2%, p-value = 0.0838) or gender (female: 46.4% vs male: 42.9%, p-value = 0.7258). At baseline, mean HDRS score was significantly higher for the chronic depressive class compared to the responder class (mean (sd): 20.9 (6.1) vs 17.9 (5.4), p-value = 0.0180). There were no significant differences between the two groups in hospitalization for CHF, cardiovascular history or the other physical health conditions examined. Although the chronic depressive group reported a larger percentage of cardiac procedures, the only significant cardiac difference was presence of an implantable cardioverter defibrillator (chronic depressives, 35.7% vs responders, 17.1%, p-value = 0.0191). The chronic depressives were more likely to have a history of anxiety. A larger proportion in the chronic depressive class was on calcium channel blockers (21.4% vs responders, 8.3%; p-value = 0.0288), but a larger proportion of the responder class was on aspirin (85.8% vs chronic depressive class 71.4%; p-value = 0.0501).

During the acute phase (Table 2), unstable angina was significantly more frequent in the chronic depressive class than in the responder class (mean (sd): 0.25 (.80) vs 0.04 (0.32), p -value = 0.0059), but otherwise there was no significant difference in cardiac events or status. Over the period of a year (Table 2), only unstable angina distinguished the two classes (chronic depressives, 0.39 (0.96) vs responders, 0.11 (0.42), p -value = 0.0266); 29% in each class died, while dropout ranged from 35.7% (chronic depressives) to 41.5% (responders). Death, dropout and reason for dropout did not differ between the two groups. Remission, defined alternatively as HDRS <8, and 50% or greater reduction in baseline HDRS score, occurred only in the responder group, where, at acute phase end 63% had an HDRS score <8, and for 66% score was at least 50% less than at baseline.

Placebo group

A three class piece-wise linear growth model best fit HDRS scores for the placebo group (Figure 3). The three classes were (1) a moderating depressive class ($n=44$, 18.8%) with HDRS score declining monotonically from an estimated baseline (intercept) mean of 20.3 to an HDRS score of 10.7 at week 14; (2) a responder class ($n=159$, 67.9%), with estimated baseline mean HDRS score of 18.1, which dropped to an estimated 7.7 at week 2, and 4.5 at week 14; and (3) a temporary improver class ($n=31$, 13.2%), with a baseline HDRS score of 16.6 that dropped to 7.4 at week 2, but then increased to a mean of 13.7 by week 14.

Age (Table 3), did not distinguish the moderating depressive, temporary improver, and responder classes (60.2, 60.6, 62.2 years respectively; $p=0.3093$), and neither did race, but gender did. The temporary improver class had the largest proportion of women ($n=18$, 56.1%), and the responder class the smallest ($n=49$, 30.8%; $p=0.0051$). At baseline, there was a significant difference in HDRS score (mean (standard deviation) chronic class: 20.6 (3.9), temporary improver: 16.2 (5.3), responder: 18.0 (5.5), p -value=0.0007). There was no significant difference among the three trajectory classes with respect to cardiovascular history or the other health conditions examined, with the exception of renal disease (chronic depressive [n (%)] 20 (45.5%), temporary improver: 5 (16.1%), responder: 61 (38.4%), p -value=0.0262). Neither was there any statistically significant difference in cardiac medications.

During the acute phase and one-year follow-up (Table 4), there was also no significant difference in cardiac events or survival among the three classes. Over the period of a year, death and proportion of dropouts did not differ significantly across the three classes, but the responder class was more likely to leave the study because of noncompliance. There were significant differences in the percentage remitting in each class (HDRS score < 8 at acute phase end) (moderating depressive: 21%, temporary improvers: 3%, responders: 70%), as well as in the proportion with a 50% or greater reduction in HDRS score (34%, 3%, 76% respectively).

We checked to see whether the treatment and placebo responder classes were comparable, and also whether the nonresponder classes in the two groups were comparable. The responder class in the treatment group had a significantly larger percent of females than the placebo responder group (88 (42.9%) vs 49 (30.8%) p -value = 0.0180). In the comparison of treatment and placebo nonresponders, a larger percent of treatment than placebo nonresponders took calcium channel blockers (6 (21.4%) vs 5 (6.7%) p -value=0.0376), and the treatment non-responders had more myocardial infarctions than placebo nonresponders (mean (sd): 0.14 (0.45) vs 0.01 (0.12), p -value = 0.0283).

Discussion

Random assignment is a well-accepted procedure used to equalize the characteristics of people assigned to each group. This approach, however, still ignores the fact that each group may be constituted of people representative of different latent classes, and that the trajectories of these latent classes may differ significantly. So, within a group, there may be classes of people with different responses to an intervention, which, averaged across the classes, may suggest that no effect is present [25–27]. We are cognizant of the limitations of exploratory findings based on statistically-driven predictors, and of the fact that GMM cannot override the importance of theory-driven predictions regarding the subgroups identified.

In the SADHART-CHF trial [16], treatment with sertraline seemed to provide no greater reduction in depression or improvement in cardiovascular status in severely impaired HF patients with major depressive disorder than was found in the placebo group. Using a GMM approach with the same sample, the focus of the present analysis was to determine (1) whether unique subgroups were present in the treatment and placebo groups, and (2) whether, in the treatment group, the characteristics of the subgroups differed from each other in a manner that would permit identification of persons responsive to sertraline.

GMM indicated two latent class trajectories in the treatment group, and three latent class trajectories in the placebo group. The two latent classes in the treatment group consisted of a small class (12% of the patients) that remained chronically depressed, and a larger responder class, consisting of 88% of the treatment group. On average the responder class attained an HDRS score of 5.5, which is in the non-depressed range [28]. The original analytic approach [16], which was not designed to identify distinguishable subgroups, indicated that on average, treatment with sertraline reduced level of depression from moderate to low mild. Our analysis, however, indicates that while there was distinct improvement to a non-depressed level for the overwhelming majority, the depression level of approximately one person in eight changed minimally over the acute phase of the study.

To determine whether the effect was attributable to the antidepressant, and not to other factors associated with the intervention, it is necessary to examine the findings from the placebo group. Three latent class trajectories were found in the placebo group. One class (consisting of 68% of the sample), was clearly of responders. Their HDRS scores improved from an estimated mean of 18.1 to an estimated mean of 4.5 at acute phase end, i.e., their estimated mean score declined to the non-depressed range. This finding may reflect the influence of study facilitators whose activities (in both treatment and placebo groups) may have alleviated depression directly or by encouraging medication adherence [16]. On average, both of the other two classes (moderating depressives, temporary improvers), remained depressed, but the moderating depressives improved from a moderate to a mild level of depression. Focusing on responders, according to our findings, sertraline appears to have been effective in 20% more patients than has placebo – 88% in the treatment group improved to a non-depressed level, compared to 68% in the placebo group. While 20% is small, it is not a negligible proportion, particularly given the high responsivity in the placebo group.

The presence of two latent class trajectories in the treatment group but three classes in the placebo group demands an explanation. Without further independent data we cannot say whether these findings would be replicated in another sample of patients with severe heart failure diagnosed with major depression. It is, however, possible that the small temporary improver class in the placebo group consists of chronic depressed patients who were responsive to some event or events that temporarily alleviated their depression. The nature

of the event is not known, but in future studies it may be well to inquire whether positive, possibly temporary, events occurred, and to determine whether they were associated with an improvement in depression, as measured by the HDRS. The improvement, albeit temporary, in this group suggests that their depression can be alleviated.

We had hoped to be able to identify at baseline, characteristics that would indicate which persons were likely to be responders, and which were not. While a few statistically significant differences were present (e.g., a slightly lower HDRS score, reduced likelihood of unstable angina, less anxiety, slightly different drug regimen), the differences were not of sufficient size to be clinically useful at an individual level. Inability to identify potential responders at baseline is a limitation of most studies. It does, however, seem feasible to identify characteristics of responders within four weeks of sertraline treatment, a time when temporary improvers would be reverting to chronic depression.

Interestingly, dropout from the study, in particular among those who were non-compliant, did not indicate a non-responsive outcome, on the contrary. Perhaps in some studies, dropping out is for a positive reason (patients may decide to take charge of their lives), and does not reflect badly on the study. At least one study of enrollees in memory disorders clinics has shown that those who dropped out were less impaired [29].

It is notable that in the treatment group, the responder class was the only class in which patients remitted (63% of them), and in which any patient had a 50% or greater decline in HDRS score (66% of them). This may indicate susceptibility to both drug intervention and study participation. In the placebo group, remission, defined as HDRS < 8 at end of study, and alternatively, 50% decline in HDRS score was more notable in the responder class (70% and 76% respectively), but was present also in moderating depression (21%, 34% respectively) and minimally (3%, 3%) in the temporary improver classes, indicating, perhaps, a susceptibility to the effects associated here with study participation. The current analysis suggests that, for some patients, sertraline has a greater effect than the extent of improvement indicated by the original study [16]. Rather, it suggests that a large proportion of patients – 88% -- would no longer be classified as depressed according to HDRS criteria.

Other analyses of the same data set pooled the sample (since no significant difference had been found between the treatment and placebo groups), dropped patients with no followup data, and focused on characteristics associated with remission (i.e., had a post-acute phase HDRS score < 8) [12]. The resulting sample of 402 patients indicated that, overall, 51.7% remitted with 54.1% and 49.5% in remission from the treatment and placebo groups respectively. Crude examination of baseline HDRS scores categorized as mild (HDRS score 8–17), moderate (18–22), and severe depression (>22), found an inverse association, with remission occurring in 62.1% (mild depression), 48.9% (moderate), and 35.9% (severe). The only other characteristic significantly associated with remission was hospitalization for CHF after entry into the study.

Various models to assess intervention effects are available. They include analyses focused on endpoints (which may be pre-determined cut-points, or 50% reduction in score since baseline) or mixed models approaches that use data from all available time points, and model change over time in each of the groups at interest. GMM also uses data from all time points, but it does not assume a single population within any one group. Rather, it checks whether subgroups, classified by different growth trajectories determined by intervention impact are present, and identifies the characteristics of the trajectories.

Limitations

Our data indicate few differences between the responder classes in the two groups, or between the nonresponder classes, suggesting that these classes have probably been identified correctly. We mention this cautiously, because we were also unable to identify characteristics that distinguish responders from nonresponders. It is possible that such inability is attributable to the reduced diversity in the sample. All participants had to be CHF patients with severe depression. Further, the variables that were available focused on issues related to their cardiac condition. There was restricted information on depression-related characteristics that might have been useful in distinguishing between the classes; lifetime history of depression was limited, and information on social or economic support systems was not available. Those enrolled were very sick patients. Within 12 weeks of enrolling, 11 in the treatment and 8 in the placebo group (4.1% of the total sample), dropped out due to death; and an additional 14, who dropped out for any of the other reasons, had died before the followup phase, i.e., 7% of the sample had died by acute phase end. Given the uniformly poor health status of the group, expectation of substantial, sustained, improvement should be tempered (although improvement certainly occurred), and diverse latent classes, although found, may be fewer, or have a different trajectory, than in a healthier group or one with more diverse information. In a sample with a broader range of problems it may be possible to better identify who will respond to a given intervention, and who is less likely to do so. Clearly, analysis of additional intervention trials is needed before the value of the current analytic approach can be established. We suggest, however, that this is a promising start.

Conclusions

SADHART-CHF was a well-monitored study, which complied with the standard analytical methods for evaluating health care interventions [30]. These methods may not adequately identify subgroups of drug treatment responders, and could result in a lack of recognition of an effective drug. Given the risks attendant on depression, and the difficulties in identifying an antidepressant that works, losing a possibly effective drug, even one effective for only a subgroup, would be unfortunate. At a minimum, when standard analytic approaches indicate that there is no significant difference between drug and placebo groups, alternative statistical procedures should be used to determine whether an apparent lack of effectiveness is actually present, or whether it is a statistical artifact driven by the responses of a particular subset of participants.

Acknowledgments

Funding: Supported in part by Clinical and Translational Science Award 5UL1 RR024128-04 to Duke University from NCRR/NIH, NIMH grant 2P50-MH60451 and NIA 5P30 AG028716 (Pepper OAIC).

References

1. Blazer D. Depression in late life: review and commentary. *J Gerontol A Bio Sci Med Sci*. 2003; 58A:249–265. [PubMed: 12634292]
2. Pozuelo L, Zhang J, Franco K, Tesar G, Penn M, Jiang W. Depression and heart disease: what do we know, and where are we headed? *Cleveland Clinic J Med*. 2009; 76:59–70.
3. Alexopoulos GS. Depression in the elderly. *Lancet*. 2005; 365:1961–1970. [PubMed: 15936426]
4. Bisschop MI, Kriegsman DMW, Deeg DJH, Beekman ATF, van Tilburg W. The longitudinal relation between chronic diseases and depression in older persons in the community: the Longitudinal Aging Study Amsterdam. *J Clin Epidemiol*. 2004; 57:187–194. [PubMed: 15125629]
5. Blazer DG, Hybels CF. What symptoms of depression predict mortality in community-dwelling elders? *J Am Geriatr Soc*. 2004; 52:2052–2056. [PubMed: 15571541]

6. Charlson M, Peterson JC. Medical comorbidity and late life depression: What is known and what are the unmet needs? *Biol Psychiat*. 2002; 52:226–2235. [PubMed: 12182928]
7. Huang C-Q, Dong B-R, Lu Z-C, Yue J-R, Liu Q-X. Chronic diseases and risk for depression in old age: a meta-analysis of published literature. *Ageing Res Reviews*. 2010; 9:131–141.
8. Carney RM, Freedland KE. Treatment-resistant depression and mortality after acute coronary syndrome. *Am J Psychiatry*. 2009; 166:410–417. [PubMed: 19289455]
9. Trivedi MH, Fava M, Wisniewski SR, Rush JA, Thase ME, Quitkin F, et al. Medication augmentation after the failure of SSRIs for depression. *N Eng J Med*. 2006a; 354(12):1243–1252.
10. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, Norquist G, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*. 2006b; 163:28–40. [PubMed: 16390886]
11. Papakostas GI, Fava M. Predictors, moderators, and mediators (correlates) of treatment outcome in major depressive disorder. *Dialogues Clin Neurosci*. 2008; 10:439–451. [PubMed: 19170401]
12. Jiang W, Krishnan R, Kuchibhatla M, Cuffe MS, Martsberger C, Arias RM, O'Connor CM. SADHART-CHF Investigators. Characteristics of depression remission and its relation with cardiovascular outcome among patients with chronic heart failure (from the SADHART-CHF Study). *Am J Cardiol*. 2011; 107:545–551. [PubMed: 21295172]
13. Muthén, B. Latent variable mixture modeling. In: Marcoulides, GA.; Schumacker, RE., editors. *New Developments and Techniques in Structural Equation Modeling*. Mahwah, NJ: Lawrence Erlbaum Associates; 2001a. p. 1-33.
14. Muthén, B. Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In: Collins, LM.; Sayer, A., editors. *New Methods for the Analysis of Change*. Washington, D.C: American Psychological Association; 2001b. p. 291-322.
15. Jiang W, O'Connor C, Silva SG, Kuchibhatla M, Cuffe MS, Callwood DD, et al. Safety and efficacy of sertraline for depression in patients with congestive heart failure (SADHART-CHF): a randomized, double-blind, placebo-controlled trial of sertraline for major depression with congestive heart failure. *Am Heart J*. 2008; 156:437–444. [PubMed: 18760123]
16. O'Connor CM, Jiang W, Kuchibhatla M, Silva SG, Cuffe MS, Callwood DD, et al. Safety and efficacy of sertraline for depression in patients with heart failure: Results of the SADHART-CHF trial. *J Am Coll Cardiol*. 2010; 56:692–699. [PubMed: 20723799]
17. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4. Washington, DC: American Psychiatric Association; 1994.
18. Hamilton M. Rating depressive patients. *J Clin Psychiat*. 1980; 41:21–44.
19. Muthén, B.; Muthén, L. *M-plus user's guide*. Los Angeles, CA: Muthén & Muthén; 2004.
20. Nagin, DS. *Group-Based Modeling of Development*. Harvard University Press; Cambridge, MA: 2005.
21. Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*. 1999; 55:463–469. [PubMed: 11318201]
22. Muthén B. Beyond SEM: General latent variable modeling. *Behaviormetrika*. 2002; 29:81–117.
23. Lo Y, Mendell NR, Rubin DB. Testing the number of components in a mixture. *Biometrika*. 2001; 88:767–778.
24. Nylund KL, Asparouhov T, Muthen B. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*. 2007; 14:535–569.
25. Mora PA, Bennett IM, Elo IT, Mathew L, Coyne JC, Culhane JF. Distinct trajectories of perinatal depressive symptomatology: evidence from growth mixture modeling. *Am J Epidemiol*. 2009; 169:24–32. [PubMed: 19001135]
26. Terrera GM, Brayne C, Matthews F. One size fits all? Why we need more sophisticated analytic methods in the explanation of trajectories of cognition in older age and their potential risk factors. *Int Psychogeriatrics*. 2010; 22:291–299.
27. Wilkosz PA, Seltman HJ, Devlin B, Weamer EA. Trajectories of cognitive decline in Alzheimer's disease. *Int Psychogeriatrics*. 2010; 22:281–290.

28. Burns, A.; Lawlor, B.; Craig, S. *Assessment Scales in Old Age Psychiatry*. London, UK: Martin Dunitz Ltd; 1999. p. 4
29. Smith DS, Fillenbaum GG. Comparison of spouse and nonkin controls: The experience of the Consortium to Establish a Registry for Alzheimer's Disease. *Aging: Clin Exp Res*. 1994; 6:151–157.
30. Schulz KF, Altman DG, Moher D. for the CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *PLoS Med*. 2010; 7(3):e1000251. [PubMed: 20352064]

\$watermark-text

\$watermark-text

\$watermark-text

I, S, Q = Intercept, linear and quadratic growth factors

Figure 1.
Growth Mixture Model

Figure 2.
Trajectory Classes for Treatment Groups

Figure 3.
Trajectory Classes for Treatment Group

Table 1

Treatment Group -- Baseline Characteristics by HDRS Trajectory Class

Variable	Chronic Depressive Class (n=28, 12%)	Responder Class (n=205, 88%)	P-value
Demographic characteristics			
Age (mean (sd))	61.9 (11.7)	63.1 (10.4)	0.3940
Race (White) (N (%))	20 (71.4%)	111 (54.2%)	0.0838
Female (N (%))	13 (46.4%)	88 (42.9%)	0.7258
Hospitalizations			
No. of congestive heart failure hospitalizations in past 12 months (mean (sd) range)	1.3 (1.6) 0–4	1.2 (1.5) 0–6	0.8659
Cardiovascular History (N (%))			
Unstable angina	18 (60.9)	113 (55.1)	0.3593
Coronary artery disease	22 (78.6)	149 (72.7)	0.5084
Myocardial infarction	15 (53.6)	106 (51.7)	0.8531
Arrhythmia	14 (50.0)	79 (38.5)	0.2453
Hypertension	25 (89.3)	186 (90.7)	0.8061
Hyperlipidemia	19 (67.9)	161 (78.5)	0.2061
Diabetes	17 (60.7)	108 (52.7)	0.4241
Renal disease	11 (39.3)	69 (33.7)	0.5564
Percutaneous coronary intervention/stent	11 (39.3)	72 (35.1)	0.6661
Coronary artery bypass graft surgery	11 (39.3)	71 (34.6)	0.6288
Implantable cardioverter defibrillator	10 (35.7)	35 (17.1)	0.0191
Permanent pacemaker	3 (10.7)	14 (6.8)	0.4584
Depression and Anxiety (N (%))			
Treated for depression	3 (10.7)	16 (7.8)	0.5977
Anxiety *	3 (10.7)	4 (1.9)	0.0393
HDRS, mean (sd)	20.9 (6.1)	17.9 (5.4)	0.0180
Cardiac Medications at baseline (N (%))			
Ace inhibitor	18 (64.3)	149 (72.7)	0.3550
Beta blocker	25 (89.3)	176 (85.8)	0.6207
Calcium channel blocker	6 (21.4)	17 (8.2)	0.0288
Aspirin	20 (71.4)	176 (85.8)	0.0501
Digoxin	4(14.3)	32 (15.6)	0.8557
Statin	20 (71.4)	147 (71.7)	0.9755

* Fisher's Exact Test; HDRS = Hamilton Depression Rating Scale; sd = standard deviation

Table 2

Treatment Group -- Acute Phase and One-Year Outcome Events

Variable	Chronic depressive Class (n=28, 12%)	Responder Class (n=205, 88%)	P-value
Acute Phase Events			
Composite Cardiac Events, n (%)			
Worsened	8 (28.6)	62 (30.2)	0.4238
Unchanged	11 (39.3)	57 (27.8)	
Improved	9 (32.1)	86 (42.0)	
Improvement in NYHA class, n (%)	9 (32.1)	89 (43.4)	0.2571
Death, * n (%)	0 (0.0)	11 (5.4%)	0.3690
Unstable angina, mean (sd) range	0.25 (0.80) 0–4	0.04 (0.32) 0–4	0.0059
Myocardial infarction, mean (sd) range	0 (0) 0	0.01 (0.12) 0–2	0.5259
Arrhythmia, mean (sd) range	0.04 (0.19) 0–1	0.03 (0.19) 0–2	0.7318
Cardiovascular event, mean (sd) range	0.57 (0.96) 0–4	0.43 (0.79) 0–4	0.4372
Congestive heart failure event, mean (sd) range	0.21 (0.50) 0–2	0.20 (0.51) 0–4	0.8538
Stroke, mean (sd) range	0 (0) 0	0.03 (0.20) 0–2	0.4087
Syncope, mean (sd) range	0.0 (0.0) 0	0.01 (0.07) 0–1	0.3639
Dropouts plus deaths, n (%)	10 (35.7)	85 (41.5)	0.5615
Reason for Dropout			
Patient withdrew	0	6 (2.9)	0.3591
Lost to followup	0 (0.0)	5 (2.4)	0.4035
Side-effects	2 (7.1)	25 (12.2)	0.4334
Withdrawn	4 (14.3)	11 (5.4)	0.0712
Non-compliance	3 (10.7)	23 (11.2)	0.9365
Other	1 (3.6)	4 (2.0)	0.5789
Outcome at One Year Followup			
Death, n (%)	8 (28.6)	59 (28.8)	0.9817
Unstable angina, mean (sd) range	0.39 (0.96) (0–4)	0.11 (0.42) (0–4)	0.0266
Arrhythmia, mean (sd) range	0.25 (0.97) (0–5)	0.08 (0.38) (0–3)	0.4504
Cardiovascular event, mean (sd) range	1.89 (2.08) (0–6)	1.22 (2.33) (0–22)	0.0535
Myocardial infarction, mean (sd) range	0.14 (0.45) (0–2)	0.03 (0.18) (0–1)	0.0717
Congestive heart failure, mean (sd) range	0.64 (1.39) (0–5)	0.59 (1.50) (0–11)	0.8699
Stroke, mean (sd) range	0.03 (0.19) (0–1)	0.05 (0.24) (0–2)	0.8416

Variable	Chronic depressive Class (n=28, 12%)	Responder Class (n=205, 88%)	P-value
Syncope, mean (sd) range	0.0 (0.0) 0	0.02 (0.17) 0–1	0.3639
Remission (HDRS < 8) n (%) **	0 (0)	106 (63.1)	0.0001
50% reduction from baseline HDRS n (%) **	0 (0)	111 (66.1)	0.0001

* Fisher's Exact Test; HDRS = Hamilton Depression Rating Scale; sd = standard deviation

** Percentages based on the sample of 402 subjects with repeat HDRS

Table 3

Placebo Group -- Baseline Characteristics by HDRS Trajectory Class

Variable	Moderating Depressive Class (n=44, 18.8%)	Temporary Improver Class (n=31, 13.2%)	Responder Class (n=159, 67.9%)	p-value
Demographic characteristics				
Age (y), mean (sd)	60.2 (12.5)	60.6 (8.9)	62.2 (10.2)	0.3093
Race (White), n (%)	26 (59.1)	20 (64.5)	89 (56.0)	0.6640
Female, n (%)	21 (47.7)	18 (56.1)	49 (30.8)	0.0051
Hospitalizations				
No. of congestive heart failure hospitalizations in past 12 months, mean (sd) range	1.73 (1.56) (0–5)	1.35 (1.52) (0–5)	1.46 (1.67) (0–6)	0.3476
Cardiovascular History (N (%))				
Unstable Angina	28 (63.6)	21 (67.7)	84 (52.8)	0.1852
Coronary artery disease	29 (65.9)	18 (58.1)	111 (70.3)	0.3958
Myocardial infarction	20 (45.5)	11 (35.5)	73 (45.9)	0.5585
Arrhythmia	20 (45.5)	10 (32.3)	76 (47.8)	0.2824
Hypertension	35 (79.6)	26 (83.9)	138 (86.8)	0.4815
Hyperlipidemia	37 (84.1)	26 (83.9)	119 (74.8)	0.2904
Diabetes	23 (52.3)	15 (48.4)	70 (44.0)	0.6020
Renal disease	20 (45.5)	5 (16.1)	61 (38.4)	0.0262
Percutaneous coronary intervention/stent	19 (43.2)	12 (38.7)	56 (35.2)	0.6153
Coronary artery bypass graft surgery	17 (38.6)	8(25.8)	53(33.3)	0.5099
Implantable cardioverter defibrillator (ICD)	11 (25.0)	5 (16.1)	29 (18.2)	0.5391
Permanent pacemaker	1 (2.3)	1 (3.23)	10 (6.3)	0.4945
Depression and Anxiety (N (%))				
Treated for depression	3 (6.8)	3 (9.7)	9 (5.7)	0.7002
Anxiety *	1 (2.3)	1 (3.2)	4 (2.5)	0.9652
HDRS score	20.6 (3.9)	16.2 (5.3)	18.0 (5.5)	0.0007
Cardiac Medications at Baseline (N (%))				
Ace inhibitor	32 (72.7)	23 (74.2)	109 (68.6)	0.7505
Beta blocker	35 (79.6)	26 (83.9)	132 (83.0)	0.8455
Calcium channel blocker	5 (11.4)	0 (0)	13 (8.2)	0.1763
Aspirin	38 (86.4)	25 (80.7)	130 (81.8)	0.7452
Digoxin	8 (18.2)	5(16.1)	31 (19.5)	0.9019
Statin	34 (77.3)	19 (61.3)	104 (65.4)	0.2538

* Fisher's Exact Test; HDRS = Hamilton Depression Rating Scale; sd = standard deviation

Table 4

Placebo Group -- Acute Phase and One-Year Outcome Events

Variable	Moderating Depressive Class (n=44, 18.8%)	Temporary Improver Class (n=31, 13.2%)	Responder Class (n=159, 67.9%)	p-value
Outcomes During Acute Phase				
Composite Cardiac Events, n (%)				0.4217
Worsened	14 (31.8)	11 (35.5)	47 (29.6)	
Unchanged	9 (20.5)	11 (35.5)	39 (24.5)	
Improved	21 (47.7)	9 (29.1)	73 (45.9)	
Improvement in NYHA class, n (%)	21 (47.7)	10 (32.3)	74 (46.5)	0.3138
Death (n (%))	2 (4.6)	1 (3.2)	4 (2.5)	0.780
Unstable angina, mean (sd) range	0.02 (.15) (0–1)	0.10 (.30) (0–1)	0.04(.27) (0–3)	0.1242
Myocardial infarction, mean (sd) range	0(0) 0	0.03 (0.18) 0–1	0.01 (0.11) 0–1	0.4745
Arrhythmia, mean (sd) range	0.07 (0.25) 0–1	0(0) 0	0.04 (0.30) 0–3	0.2008
Cardiovascular event, mean (sd) range	0.52 (1.02) 1–5	0.39 (0.56) 0–2	0.49 (1.02) 0–6	0.7596
Congestive heart failure, mean (sd) range	0.34 (0.89) 0–5	0.16 (0.37) 0–1	0.25 (0.69) 0–5	0.7421
Stroke, mean(sd) range	0(0) 0	0(0) 0	0.01 (0.08) 0–1	0.7899
Syncope, mean (sd) range	0(0) 0	0 (0) 0	0.01 (.08) 0–1	0.7899
Dropouts plus deaths n (%)	13 (29.6)	7 (22.6)	62 (39.0)	0.1504
Reason for dropout, n (%)				
Patient withdrew	2 (4.6)	0 (0.0)	6 (3.8)	0.6419
Lost to followup	2 (4.6)	1 (3.2)	6 (3.8)	0.8774
Side-effects	3 (6.8)	0 (0.0)	11 (6.9)	0.3611
Withdrawn	3 (6.8)	3 (9.7)	9 (5.7)	0.5500
Noncompliance	1 (2.3)	0 (0.0)	23 (14.5)	0.0051
Outcomes at One Year Follow-up				
Death, n (%)	13 (29.6%)	5 (16.1%)	42 (26.4%)	0.3920
Unstable angina, mean (sd) range	0.14 (0.41) 0–2	0.32 (1.01) 0–5	0.08 (0.36) 0–3	0.2917
Myocardial infarction, mean (sd) range	0 (0) 0	0.03 (0.18) (0–1)	0.05 (0.25) (0–2)	0.3640
Arrhythmia, mean (sd) range	0.14 (0.41) 0–2	0(0) 0	0.14 (0.50) 0–3	0.1783
Cardiovascular, mean (sd) range	1.27 (2.05) 0–11	1.29 (2.30) (0–10)	1.30 (1.96) (0–13)	0.8361
Congestive heart failure, mean (sd) range	0.72 (1.50) 0–7	0.32 (0.75) 0–3	0.64 (1.38) 0–11	0.4113
Stroke, mean (sd) range	0.04 (0.30) 0–2	0(0)	0.03 (0.24) 0–2	0.7221

Variable	Moderating Depressive Class (n=44, 18.8%)	Temporary Improver Class (n=31, 13.2%)	Responder Class (n=159, 67.9%)	p-value
Syncope, mean (sd) range	0.03 (0.24) 0–1	0.06 (.25) 0–1	0.02 (0.14) 0–1	0.3374
Remission, n (%) ** (HDRS < 8)	9 (20.5)	1 (3.2)	92 (70.2)	0.0001
50% reduction in HDRS from baseline, n (%) **	15 (34.1)	1 (3.2)	99 (75.6)	0.0001

* Fisher's Exact Test; HDRS = Hamilton Depression Rating Scale; sd = standard deviation

** Percentages based on the sample of 402 subjects with repeat HDRS