Prognostic Subgroups for Citalopram Response in the STAR*D Trial

Ewgeni Jakubovski, MA and
Connecticut Mental Health Center and the Yale Child Study Center, Yale University, New Haven, Connecticut

Michael H. Bloch, MD, MS
Child Study Center and Department of Psychiatry, Yale University, New Haven, Connecticut

Abstract

Objective—Few data exist to help clinicians predict likelihood of treatment response in individual patients with major depressive disorder (MDD). Our aim was to identify subgroups of MDD patients with differential treatment outcomes based on presenting clinical characteristics. We also sought to quantify the likelihood of treatment success based on the degree of improvement and side effects after 2 and 4 weeks of selective serotonin reuptake inhibitor (SSRI) pharmacotherapy.

Method—We analyzed data from the first treatment phase of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial, in which subjects with a *DSM-IV* diagnosis of MDD were treated for 8–14 weeks with open-label citalopram. A receiver operating characteristic (ROC) analysis was conducted to determine homogenous subgroups with different rates of response and remission in depressive symptoms. Included predictor variables were initial clinical characteristics, initial improvement, and side effects after 2 and 4 weeks of SSRI treatment. The primary outcome measures were treatment response (defined as a greater than 50% reduction in 17-item Hamilton Depression Rating Scale [HDRS-17] score from baseline) and remission (defined as an HDRS-17 score 17).

Results—Baseline clinical characteristics were able to identify subgroups from a low likelihood of response of 18% (income < $ 10,000, comorbid generalized anxiety disorder, < 16 years of education; *P* <.01) to a high likelihood of response of 68% (income $40,000, no comorbid posttraumatic stress disorder; *P* <.01). Among baseline clinical characteristics, employment status (N = 2,477; X^2_1 = 78.1; *P* <.001) and income level (N = 2,512; X^2_1 =77.7; *P*< .001) were the most informative in predicting treatment outcome. For the models at weeks 2 and 4, treatment success was best predicted by early symptom improvement.

Corresponding author: Ewgeni Jakubovski, MS, Child Study Center, Yale University School of Medicine, PO Box 2070900, New Haven, CT 06520, (ewgeni.jakubovski@yale.edu).

Conclusions—Socioeconomic data such as low income, education, and unemployment were most discriminative in predicting a poor response to citalopram, even with disparities in access to care accounted for. This finding implies that socioeconomic factors may be more useful predictors of medication response than traditional psychiatric diagnoses or past treatment history.

Trial Registration—ClinicalTrials.gov identifier: NCT00021528

Antidepressant medications are the most commonly prescribed intervention for the treatment of major depressive disorder (MDD).[1–3] Conventional wisdom is that the effects of antidepressants take at least 2–4 weeks to become clinically evident, although there is evidence that the treatment effects of selective serotonin reuptake inhibitors (SSRIs) can be observed statistically as early as 1 week after the initiation of treatment.[4] Improvement with antidepressant agents continues 6, 8, and even 12 weeks following initiation of SSRI medication.[4] Approximately half of patients with MDD respond (achieve a clinically meaningful reduction in symptoms) with SSRI treatment for depression.[5] A much smaller proportion of MDD patients, 25%–40%, experience remission, the virtual absence of symptoms, after an acute antidepressant trial.[5]

A large fraction of patients do not respond or remit with SSRI treatment for MDD, which often takes several months to reach maximal benefits in reducing depressive symptoms. Therefore, identifying predictors and moderators of SSRI treatment effects is clinically important. Particularly useful, in this regard, is identifying subgroups of MDD patients with differential likelihood of improvement with SSRI pharmacotherapy. This knowledge would provide a better foundation for clinicians to make decisions at an earlier stage of treatment regarding treatment strategies for individual patients (continuation of current medication, augmenting current pharmacotherapy, or switching).[6–7]

Socioeconomic status (SES) plays an important role in the prevalence of mental disorders. According to a cross-national comparison of prevalences and correlates of mental disorders carried out by the World Health Organization, people with low SES show higher prevalence rates in almost all mental disorders.[8] Whereas poverty and unemployment were shown to be predictive of persistence of mental disorders, financial strain was associated with both onset and future morbidity.[9] As for depression in particular, a meta-analysis showed that low SES was associated with both the new onset and persistence of depression.[10] Access to mental health services varies as a function of SES, to the disadvantage of those with lower SES.[11] This association poses the question as to whether the relationship between SES and psychiatric disorder—MDD in particular—is mediated by the accessibility of treatment. This question can be answered by examining a large clinical trial providing standardized care to a sample of patients with large sociodemographic differences. Such a trial is the Sequenced Treatment Alternatives to Relieve Depression (STAR*D).[12,13]

Logistic regression, which is usually the method of choice in treatment studies,[5] is not helpful in identifying homogenous subgroups of patients with differential likelihood of responding. Its possibilities to account for potential interactions between predictor variables are strictly limited to a few a priori-defined ones. An alternative method of analysis is offered by the receiver operating characteristic (ROC) analysis, which identifies subgroups

of subjects with meaningful differences in the outcome variable from a set of predictor variables.

We applied ROC analysis to data from the first treatment phase of the STAR*D trial, which enrolled over 2,000 nonpsychotic MDD patients and treated them for 8–14 weeks with citalopram in real-world settings.[12] We sought to empirically identify homogeneous subgroups of patients with different prognoses after citalopram treatment (both response and remission) using baseline demographic, social, and clinical characteristics. We refined these models using clinical data from weeks 2 and 4 of treatment to examine the additional predictive value of information on early response and side effects to help guide clinical decisions early in SSRI treatment. These empirically derived models should help improve the accuracy of initial prognosis for individual patients treated with SSRIs and provide additional prognostic information to guide treatment decisions early in pharmacotherapy.

METHOD

Study Overview

The rationale, design, and methods of the STAR*D trial have been described in depth elsewhere.[12,13] We specifically utilized data from the first treatment phase of the STAR*D trial, which was a large, uncontrolled, practical clinical trial in which more than 2,500 subjects were treated with citalopram for 8–14 weeks. The research protocol was approved by the relevant review boards and all subjects provided informed consent. The study was registered on ClinicalTrials.gov (identifier: NCT00021528).

Subjects

Subjects were recruited from 18 primary care centers and 23 psychiatric clinical sites throughout the United States. To maximize the generalizability of the study findings to real-world settings, subjects were recruited (1) only from patients seeking medical care in routine medical or psychiatric outpatient treatment (as opposed to through advertisement), (2) through both public and private health care settings, and (3) with minimal exclusion and broad inclusion criteria.

To be included in the STAR*D trial, outpatients needed to be adults aged 18–75 years and present with a nonpsychotic major depressive episode. They were required to have a score at baseline of greater than or equal to 14 on the 17-item Hamilton Depression Rating Scale (HDRS-17).[14–15] Patients were excluded from the STAR*D trial if they were pregnant or breast-feeding or had a primary psychiatric diagnosis of bipolar disorder, a psychotic disorder, obsessive-compulsive disorder, or an eating disorder. Subjects were also excluded if they had a general medical condition that was a contraindication for the use of any antidepressant agent used in the first 2 treatment phases of STAR*D or if they had a clear history of nonresponse or intolerance to these agents.

Assessment

As per the STAR*D treatment protocol, a checklist based on *DSM-IV* criteria was used to confirm the diagnosis of nonpsychotic MDD. At baseline, self-reports were obtained to

provide information on demographic characteristics, past treatment history, and family history of Axis I psychiatric disorders. The Psychiatric Diagnostic Screening Questionnaire[16–18] was used to establish the presence of 11 potential comorbid Axis I psychiatric diagnoses. The 16-item Quick Inventory of Depressive Symptomatology, Self-Report (QIDS-SR),[19–21] and the HDRS-17 were utilized to assess depression symptom severity. The QIDS-SR ratings and side effects ratings were assessed at every subsequent treatment visit. Clinical visits were suggested at 2, 4, 6, 9, and 12 weeks of citalopram treatment during the first-phase of the STAR*D trial. Side effects were rated in terms of their frequency, intensity, and burden on 7-point Likert subscales.[12] The HDRS-17 was the identified primary outcome measure for the trial. The HDRS-17 ratings were collected by research outcome assessors via telephone-based structured interviews either in English or in Spanish at baseline and prior to exit from phase I of the STAR*D trial.

Intervention

Citalopram was prescribed in an open-label, unblinded manner to all subjects enrolled in the STAR*D protocol. The starting dose of citalopram was 20 mg per day, which was increased to 40 mg per day by week 4 and a maximum dose of 60 mg per day by week 6. However, the treatment protocol allowed for individualized starting doses and dose adjustments in order to minimize side effects, maximize safety, and optimize chances of therapeutic benefit in individual patients. Medication management was informed by QIDS ratings conducted at each study visit. The STAR*D protocol recommended treatment visits at weeks 2, 4, 6, 9, and 12. Patients were allowed to discontinue citalopram before 12 weeks if (1) they had intolerable side effects, (2) an optimal dose was not possible due to side effects or patient choice, or (3) significant depressive symptoms (Quick Inventory of Depressive Symptomatology, Clinician Rating score 9) were present after 9 weeks of treatment with citalopram at the maximum tolerated dose.

Statistical Analysis

Data preparation was conducted using SAS version 9.3 and Microsoft Excel (SAS Institute; Cary, North Carolina). Signal detection methodology was used to find the best prediction model. The ROC analysis was performed using a free software available online from Ruth O'Hara, PhD, at Stanford University (http://www.stanford.edu/~yesavage/ROC.html). Data utilized in this study were obtained from the National Institute of Mental Health-supported STAR*D Limited Access Dataset, version 2. The ROC analysis is a nonparametric method that operates via recursive partitioning. It aims to identify subgroups of individuals who have a higher or lower probability of achieving a particular binary outcome.[22] The ROC analysis has several advantages over traditional regression analyses: it has improved power and flexibility when examining higher-order interactions—ROC analysis, in contrast to regression, can analyze all possible interactions, rather than only those specified a priori, and can analyze interactions even when the main-effect terms are not included in the model. The ROC analysis can also handle missing data without discarding all other prognostic data. More information on the ROC analysis can be found elsewhere.[22–24] Remission and response at week 12 of the citalopram treatment phase were utilized as the binary outcome for ROC analysis. *Response* was defined as a greater than 50% reduction in HDRS-17 score from baseline. *Remission* was defined as an HDRS-17 score 7. We decided to look at both

remission and response as treatment outcomes because they measure related but ultimately different treatment outcomes, both of which are of significant importance to clinicians. Response identifies those individuals that improve significantly with treatment, and remission captures those individuals who have minimal symptoms after treatment. Since both remission and response are of clinical importance, we decided to present both as outcomes of the model rather than chose 1 arbitrarily. For each measured potential predictor, cutoff points are generated at all values observed in the variable. The quality of a cutoff point is defined by its ability of dividing the sample into 2 subsamples maximally distinct in the specified binary outcome. The cutoff point that yields the best prediction is identified across all values of all variables. That cutoff point is then used to divide the total sample in 2 subsamples. The same procedure is repeated systematically in each of the 2 subsamples. The sensitivity of the cutoff point was set to 0.5—a neutral value, neither conservative nor lax. This iterative process continues until a stopping criterion is reached. The following stopping criteria were applied: a subgroup size of less than 20 individuals or a failure to reach a significant group difference at $P <.05$ for any candidate cutoff value. The iterations were also stopped once the 3-way interaction level was reached. After the last step of the ROC analysis was reached, we calculated the probability of response/remission for each subgroup and presented results as hierarchical decision tree diagrams.

In total, 6 ROC analyses were conducted. Models were calculated using both response and remission as the outcome variable. Models were conducted based on clinical information that would be potentially available at baseline and then after 2 and 4 weeks of citalopram treatment. Several predictors were entered into the model:

1. Demographic predictors: age, race (white and African American [yes/no]), ethnicity (Hispanic vs non-Hispanic), and gender

2. Socioeconomic predictors: years of education, academic degrees (high school dropout, high school diploma, some college and/or graduate school), employment status (employed/unemployed), income (categorized as < $10,000; $10,000– 20,000; $20,000–$40,000; and > $40,000), marital status (never married, divorced, widowed, cohabitant, separated, and married)

3. Clinical predictors: age at onset of MDD, length of the MDD history, family history of depression, history of resistance to SSRI treatment, history of resistance to antidepressant treatment, past suicide attempts, and comorbid Axis I psychiatric disorders (posttraumatic stress disorder [PTSD], bulimia, panic disorder, agoraphobia, social phobia, alcohol dependence, drug dependence, generalized anxiety disorder [GAD], somatization disorder, obsessive-compulsive disorder, hypochondriasis).

The week 2 model additionally contained side effect variables (intensity, frequency, burden) at 2 weeks, percentage of improvement on QIDS (categorized as worsened symptoms, improvement 0%–10%, improvement 10%–20%, improvement 20%–33%, improvement 33%–49%) and remission or response at 2 weeks. The week 4 model contained all variables included in the week 2 model and, additionally, side effects, percentage of improvement on QIDS, and remission or response at 4 weeks. To look at remission and response

independently, we excluded subjects who attained remission by week 2 from the week 2 response model and vice versa. The same was done for the week 4 models. Since initial responders/remitters were excluded from the week 2 and week 4 models, these models are useful for predicting outcome only in delayed responders (eg, subjects who have not already obtained the outcome). All analyses are completer analyses; only subjects who had an assessment at 12 weeks were included.

RESULTS

Subjects

Demographic and clinical characteristics for the subjects included in each model are summarized in Table 1. The sample size for models varied from 2,512 subjects for baseline models to 991 subjects for the week 4 predictors of outcome. The decreasing sample size in later models is due to the fact that a large proportion of subjects had either responded/ remitted or dropped out by the later assessment points.

Empirically Derived Prognostic Subgroups at Baseline Associated With Treatment Outcome

Figure 1A displays empirically derived hierarchical prognostic subgroups for remission with citalopram treatment. Baseline clinical characteristics were able to identify subgroups with as low as a 15% likelihood of remitting with citalopram treatment (unemployed, low education level [< 14 years], and comorbid social phobia) to as high as a 62% likelihood of remitting with citalopram treatment (employed, high income [> $40,000], and female). The most discriminative predictor of remission was employment status (N = 2,477; χ^2_1 = 78.1; P <.001).

Figure 1B displays empirically derived hierarchical prognostic subgroups for response after citalopram treatment. Baseline clinical characteristics were able to identify subgroups with as low as an 18% likelihood of responding to citalopram (income less than $10,000, comorbid GAD, and less than 16 years of education) to as high as a 68% likelihood of responding to citalopram (income $40,000 with no comorbid PTSD). The most discriminative predictor of citalopram response was low income (threshold of $10,000: N=2,512; χ^2_1 = 77.7; P <.001). The significance level for all subsequent nodes in decision trees are displayed by dark and light gray and by white.

Empirically Derived Prognostic Subgroups After 2 Weeks of Citalopram Treatment

Figure 2 displays empirically derived prognostic subgroups for response and remission after inclusion of clinical information available after 2 weeks of citalopram treatment. The most discriminative predictor for both models was the degree of improvement in depressive symptoms measured on the QIDS for response (N= 1,641; χ^2_1 =49.9; P < .001) and for remission (N= 1,857; χ^2_1 = 66.1; P <.001). The threshold for cutoff on the QIDS was 20% improvement for the clinical response model and 33% for the clinical remission model. The clinical response model was able to identify subgroups with as low as a 26% likelihood of responding (low initial improvement in depressive symptoms [QIDS reduction <20%], low income [< $20,000], and aged 33 years) to as high as 68% (high initial improvement in

depressive symptoms [QIDS reduction   20%], low side effect intensity [rated as mild or less] and not low income [$10,000]). The remission model was able to identify subgroups with as low as 22% (small initial depression improvement [QIDS improvement < 33%], low income [< $20,000], and aged   33 years) and as high as 71% likelihood of remission (large initial depression improvement [QIDS improvement   33%], better educated [more than 14 years of education], and high income [$40,000]) based on clinical characteristics available at week 2.

Empirically Derived Prognostic Subgroups After 4 Weeks of Citalopram Treatment

Figure 3 displays empirically derived prognostic subgroups for response and remission after inclusion of clinical information available after 4 weeks of citalopram treatment. The most discriminative predictor for both remission and response models was improvement in depressive symptoms (QIDS percent improvement of 33% at week 4) for response (N = 991; $\chi^2_1 = 51.1$; $P < .001$) and for remission (N= 1,262; $\chi^2_1 = 61.4$; $P < .001$). The clinical response model was able to identify subgroups with as low as a 23% likelihood of responding (low initial improvement in depressive symptoms [QIDS reduction <20%] and low income [< $20,000]) to as high as 73% (large improvement in depressive symptoms [QIDS reduction   33%] with trivial side effect intensity and frequency). The remission model was able to identify subgroups with as low as 16% (small initial depression improvement [QIDS improvement < 20%] and not high income [< $40,000]) and as high as 61 % likelihood of remission (large initial depression improvement [QIDS improvement 33%], at least medium income [$20,000], and female), based on clinical characteristics available at week 4.

More detailed statistical information such as sensitivity, specificity, and P values for all models are provided in Supplementary eTable 1 (available at PSYCHIATRIST.COM).

DISCUSSION

The ROC analysis was able to empirically derive prognostic subgroups in MDD patients with low (eg, < 25% likelihood of responding or remitting) or high (eg, greater than 67% likelihood of responding or remitting) with citalopram treatment in the STAR*D trial. The ROC analysis was also able to identify significant higher-order interactions between baseline clinical variables that would not be evident with traditional regression analysis. For instance, in the ROC analysis of remission at baseline, a significant interaction was observed between unemployment and low education level that greatly decreased likelihood of remission.

Socioeconomic measures, such as income, employment status, and education, were the best predictors of treatment response and more discriminative than clinical attributes, such as past medication response, severity and duration of depression, comorbid psychiatric diagnoses, and substance use. These results make sense in the light of previous studies,[10,25,26] which found associations between low income and low levels of education with higher risk of depression, longer depressive episodes, and poor treatment response. Comorbid anxiety disorders (eg, GAD, social phobia, and PTSD) were the only baseline diagnoses always associated with poor outcomes, which is consistent with a previous

analysis[27] of STAR*D comparing anxious to nonanxious depression. Socioeconomic status along with measures of anxiety were also shown to be independently associated with treatment outcome in an elderly population.[28] The nature of the predictive hierarchies changed rapidly when information from early response to treatment was included. Clinical rating of the initial response to citalopram as early as 2 weeks after the start of pharmacotherapy was the most discriminative variable in predicting ultimate response or remission, outweighing baseline clinical characteristics in predictive value. Increased side effect burden early in treatment, in contrast, was associated with poor treatment outcomes across models. The cutoff point for tolerance of side effects was universally higher in the subjects experiencing greater symptom relief with citalopram treatment.

Perhaps the most striking factor regarding the association between low socioeconomic status and poor treatment outcome in STAR*D is the fact that access to treatment, quality of care, and degree of monitoring were kept quite consistent as part of the study design. Therefore, this association suggests that the association between low income, poor education, and unemployment and poor treatment outcome is independent of the quality of care received. This suggests that low SES individuals with depression are less responsive to initial pharmacologic treatment, even if disparities in access to and quality of care are equalized. The relationship between depression and SES is further supported by studies showing a positive association between SES and white matter tract integrity in the brain—which was partially mediated by factors such as adiposity and smoking.[29] From a public health perspective, more resources—increased treatment lengths, greater number of therapies received—may be needed for low SES populations with depression. Given that SES variables are most strongly associated with poor treatment outcome compared to any of the factors we examined, it may not be a realistic expectation for Medicaid or the community mental health organizations that treat these depressed patients to function as efficiently as systems treating privately insured, employed individuals. Understanding the underlying basis for the association between low socioeconomic status and poor treatment outcome in depression, independent of disparities in care, is an important research question. Socioeconomic status has been demonstrated to be a marker of underlying psychopathology; there is a downward drift of more severely affected psychiatric patients on the socioeconomic status.[30] Therefore, it remains unclear whether improving the socioeconomic status independent of treating underlying depression pathology would improve outcomes.

For clinicians, the finding that socioeconomic variables (income and employment status) are particularly discriminative in predicting treatment response highlights the importance of baseline assessment of socioeconomic variables at intake. Our analysis suggests, perhaps surprisingly, that these variables are likely to be more informative than routine clinical variables such as past medication response, duration and severity of illness, and comorbid psychiatric illnesses. Nonetheless, the ROC analysis also demonstrates on several occasions that the combination of a poor socioeconomic situation and poor clinical factors appears particularly pernicious. It should be noted that the SES variables income, length of education, and employment status show relatively low correlations in our sample ($r = 0.17$–0.34); nevertheless, all of them predict outcome in a similar way. The low correlations can be understood in terms of each variable capturing another aspect of SES.

It is perhaps less surprising, but no less important, that the predictors of response change dramatically after just 2 weeks of treatment. As early as week 2, initial treatment response was the most discriminative data point regarding likelihood of citalopram response or remission. This finding is consistent with the results of a previous meta-analysis[4] that suggests that the benefits of SSRI pharmacotherapy can be observed as early as the first week of treatment. For clinicians, these decision trees thus highlight the importance of measuring early treatment response and side effects.

Given the implications of this study, it is important to be explicit about several limitations to the current analysis. The first treatment phase of STAR*D was uncontrolled; therefore, we are unable to determine if the empirically derived prognostic subgroups generated are specific to citalopram treatment or are related to depression improvement in general. On the other hand, uncontrolled, unblinded treatment with citalopram closely resembles how this medication is used in real-world settings. This "limitation," therefore, also increases the external validity of the results, which was one of the major goals of the STAR*D trial. Moreover, the STAR*D trial was shown to be highly similar in its effectiveness to daily practice as opposed to other randomized controlled trials.[31] Another limitation is that the prognostic subgroups were empirically derived and not hypothesis driven; thus, replication of our findings is needed. However, the statistical advantages of ROC analysis allowed the exploration of higher-order interactions between clinical variables and the identification of homogenous prognostic subgroups based on easily measurable clinical characteristics. These attributes should make our analysis useful in providing prognostic information about citalopram response in clinical settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Cuijpers P, van Straten A, Schuurmans J, et al. Psychotherapy for chronic major depression and dysthymia: a meta analysis. http://dare.ubvu.vu.nl/handle/1871/16585. Updated 2009.

2. Cuijpers P, van Straten A, van Oppen P, et al. Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? a meta-analysis of comparative studies. Clin Psychiatry. 2008; 69(11):1675–1685. quiz 1839–1841.

3. Marcus SC, Olfson M. National trends in the treatment for depression from 1998 to 2007. Arch Gen Psychiatry. 2010; 67(12):1265–1273. [PubMed: 21135326]

4. 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. [PubMed: 17088502]

5. Trivedi MH, Rush AJ, Wisniewski SR, et al. STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry. 2006; 163(1):28–40. [PubMed: 16390886]

6. Mulsant BH, Houck PR, Gildengers AG, et al. What is the optimal duration of a short-term antidepressant trial when treating geriatric depression? J Clin Psychopharmacol. 2006; 26(2):113–120. [PubMed: 16633138]

7. Andreescu C, Lenze EJ, Dew MA, et al. Effect of comorbid anxiety on treatment response and relapse risk in late-life depression: controlled study. Br J Psychiatry. 2007; 190(4):344–349. [PubMed: 17401042]

8. WHO International Consortium in Psychiatric Epidemiology. Cross-national comparisons of the prevalences and correlates of mental disorders. Bull World Health Organ. 2000; 78(4):413–426. [PubMed: 10885160]

9. Weich S, Lewis G. Poverty, unemployment, and common mental disorders: population based cohort study. BMJ. 1998; 317(7151):115–119. [PubMed: 9657786]

10. Lorant V, Deliège D, Eaton W, et al. Socioeconomic inequalities in depression: a meta-analysis. Am J Epidemiol. 2003; 157(2):98–112. [PubMed: 12522017]

11. Katz SJ, Kessler RC, Frank RG, et al. Mental health care use, morbidity, and socioeconomic status in the United States and Ontario. Inquiry. 1997; 34(1):38–49. [PubMed: 9146506]

12. Rush AJ, Fava M, Wisniewski SR, et al. STAR*D Investigators Group. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. Control Clin Trials. 2004; 25(1):119–142. [PubMed: 15061154]

13. Fava M, Rush AJ, Trivedi MH, et al. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. Psychiatr Clin North Am. 2003; 26(2):457–494. x. [PubMed: 12778843]

14. Hamilton M. A rating scale for depression. J Neural Neurosurg Psychiatry. 1960; 23(1):56–62.

15. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol. 1967; 6(4):278–296. [PubMed: 6080235]

16. Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. Arch Gen Psychiatry. 2001; 58(8):787–794. [PubMed: 11483146]

17. Zimmerman M, Mattia JI. The Psychiatric Diagnostic Screening Questionnaire: development, reliability and validity. Compr Psychiatry. 2001; 42(3):175–189. [PubMed: 11349235]

18. Rush AJ, Zimmerman M, Wisniewski SR, et al. Comorbid psychiatric disorders in depressed outpatients: demographic and clinical features. J Affect Disord. 2005; 87(1):43–55. [PubMed: 15894381]

19. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry. 2003; 54(5):573–583. [PubMed: 12946886]

20. Trivedi MH, Rush AJ, Ibrahim HM, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med. 2004; 34(1):73–82. [PubMed: 14971628]

21. Rush AJ, Bernstein IH, Trivedi MH, et al. An evaluation of the quick inventory of depressive symptomatology and the Hamilton rating scale for depression: a sequenced treatment alternatives to relieve depression trial report. Biol Psychiatry. 2006; 59(6):493–501. [PubMed: 16199008]

22. Kraemer, HC. Evaluating Medical Tests: Objective and Quantitative Guidelines. Thousand Oaks, CA: SAGE Publications, Inc.; 1992. p. 296

23. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. Second Edition. New York, NY: Routledge Academic: 1988.

24. Cook EF, Goldman L. Empiric comparison of multivariate analytic techniques: advantages and disadvantages of recursive partitioning analysis. J Chronic Dis. 1984; 37(9–10):721–731. [PubMed: 6501544]

25. Ronalds C, Creed F, Stone K, et al. Outcome of anxiety and depressive disorders in primary care. Br J Psychiatry. 1997; 171(5):427–433. [PubMed: 9463600]

26. Viinamäki H, Haatainen K, Honkalampi K, et al. Which factors are important predictors of non-recovery from major depression? a 2-year prospective observational study. Nord J Psychiatry. 2006; 60(5):410–416. [PubMed: 17050300]

27. Fava M, Rush AJ, Alpert JE, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. Am J Psychiatry. 2008; 165(3):342–351. [PubMed: 18172020]

28. Cohen A, Gilman SE, Houck PR, et al. Socioeconomic status and anxiety as predictors of antidepressant treatment response and suicidal ideation in older adults. Soc Psychiatry Psychiatr Epidemiol. 2009; 44(4):272–277. [PubMed: 18818858]

29. Gianaros PJ, Marsland AL, Sheu LK, et al. Inflammatory pathways link socioeconomic inequalities to white matter architecture. Cereb Cortex. 2013; 23(9):2058–2071. [PubMed: 22772650]

30. Aro S, Aro H, Keskimäki I. Socio-economic mobility among patients with schizophrenia or major affective disorder: a 17-year retrospective follow-up. Br J Psychiatry. 1995; 166(6):759–767. [PubMed: 7663824]

31. van der Lem R, van der Wee NJA, van Veen T, et al. Efficacy versus effectiveness: a direct comparison of the outcome of treatment for mild to moderate depression in randomized controlled trials and daily practice. Psychother Psychosom. 2012; 81(4):226–234. [PubMed: 22584117]

Clinical Points

• Homogenous subgroups with different levels of response and remission with citalopram treatment are identifiable based on easily measurable baseline characteristics.

• Socioeconomic factors such as income, education, and employment status appear most discriminative in predicting citalopram response and remission.

• Within 2 weeks of starting citalopram, initial response to treatment becomes the most discriminative predictor of short-term response and remission after 12 weeks.

Figure 1.

ROC Analysis: Empirically Derived Subgroups Predicting Treatment Outcome at Week 12
Using Baseline Clinical Characteristicsa

aChart color: dark gray: *P* <.001; light gray: *P* <.01; white: *P* <.05.

Abbreviations: GAD = generalized anxiety disorder, HDRS = Hamilton Depression Rating
Scale, MDD = major depressive disorder, ROC = receiver operating characteristic.

Figure 2.

ROC Analysis: Empirically Derived Subgroups Predicting Treatment Outcome at Week 12
Using Clinical Characteristics Known After 2 Weeks of Citalopram Treatmenta
[a]Chart color: dark gray: *P* <.001; light gray: *P* <.01; white: *P* <.05.
Abbreviations: HDRS = Hamilton Depression Rating Scale, MDD = major depressive
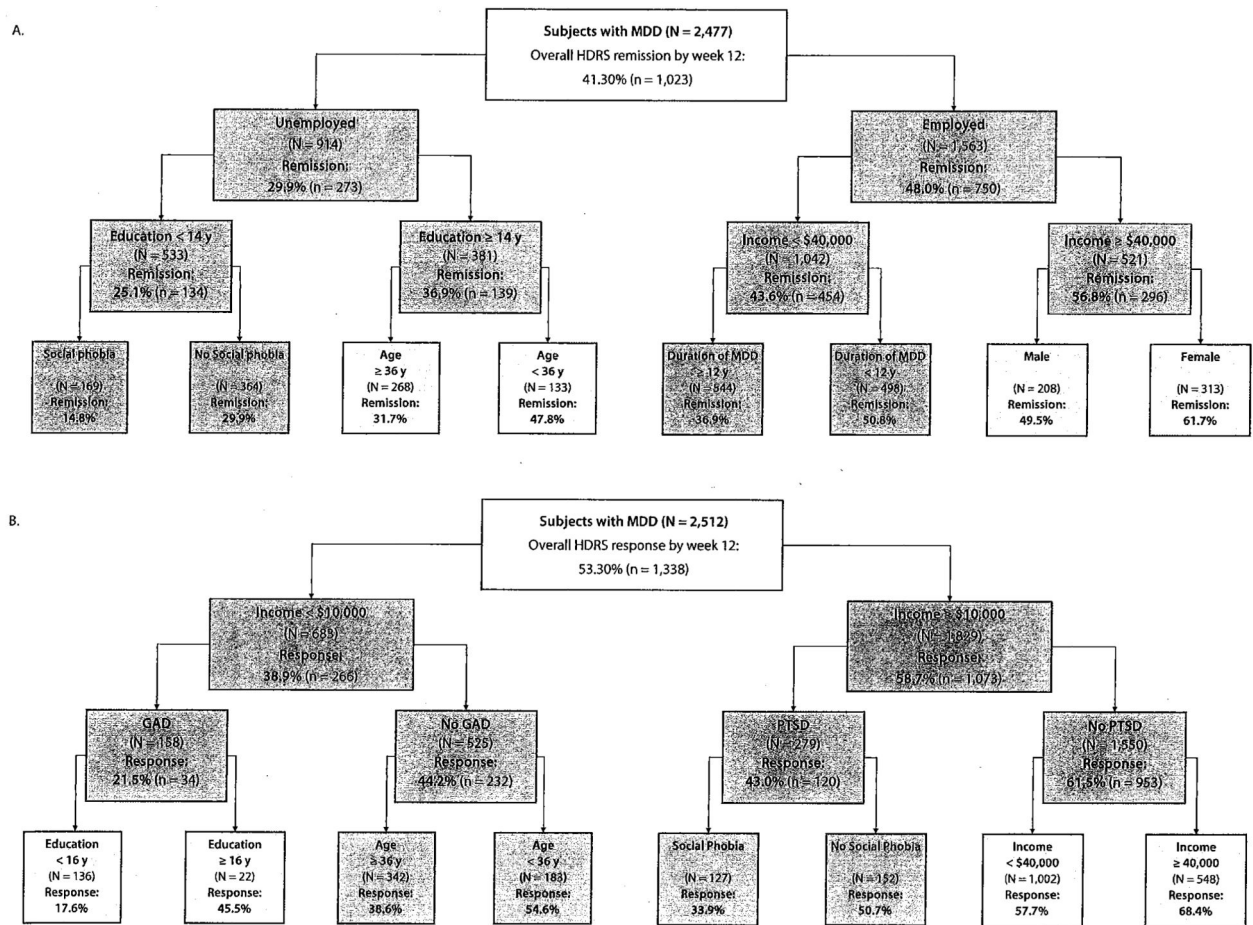disorder, QIDS = Quick Inventory of Depressive Symptomatology, ROC = receiver
operating characteristic.

Figure 3.
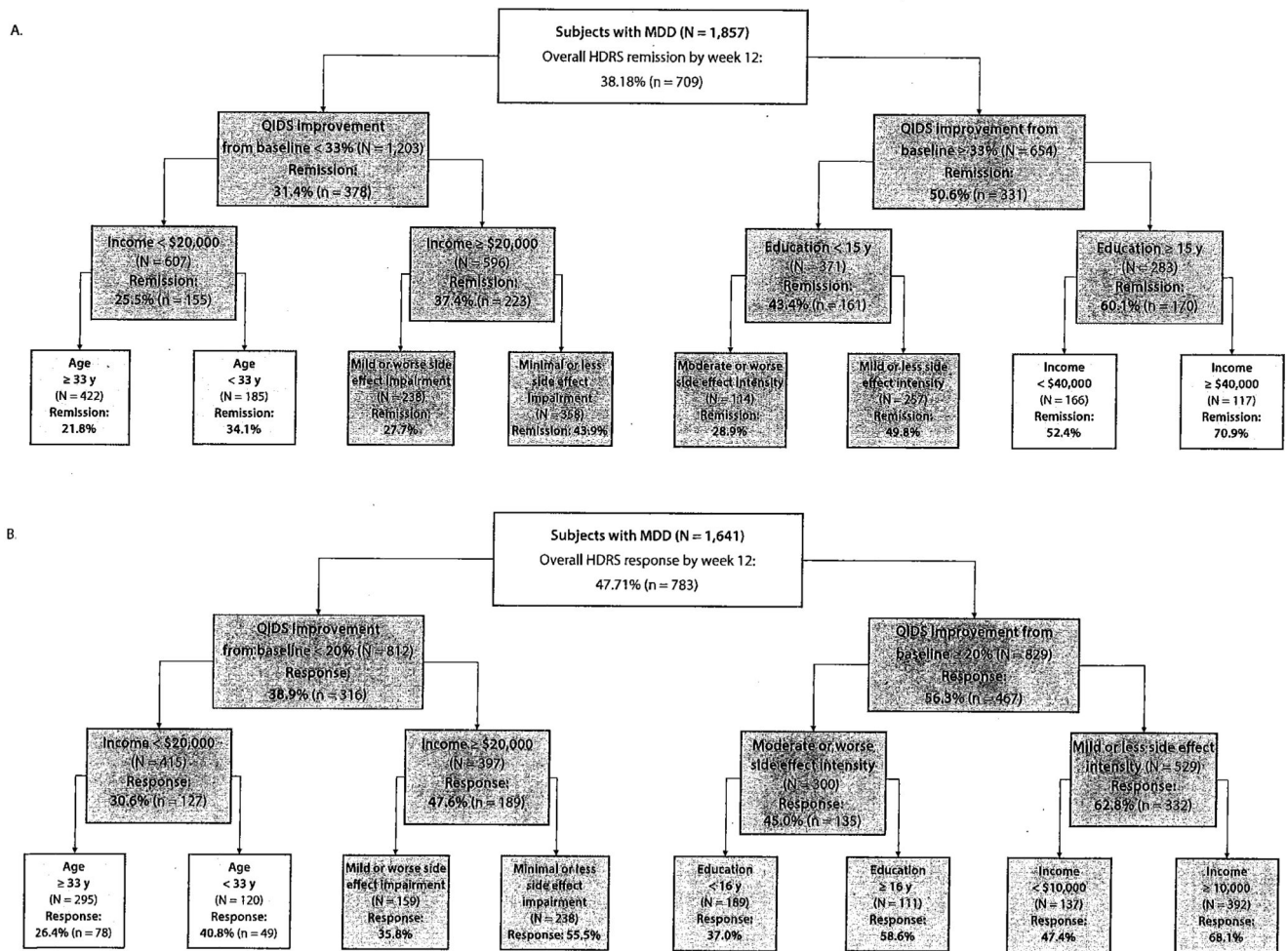ROC Analysis: Empirically Derived Subgroups Predicting Treatment Outcome at Week 12
Using Clinical Characteristics Known After 4 Weeks of Citalopram Treatment[a]
[a]Chart color: dark gray: *P* <.001; light gray: *P* <.01; White *P* <.05.
Abbreviations: HDRS = Hamilton Depression Rating Scale, MDD = major depressive
disorder, QIDS = Quick Inventory of Depressive Symptomatology, ROC = receiver
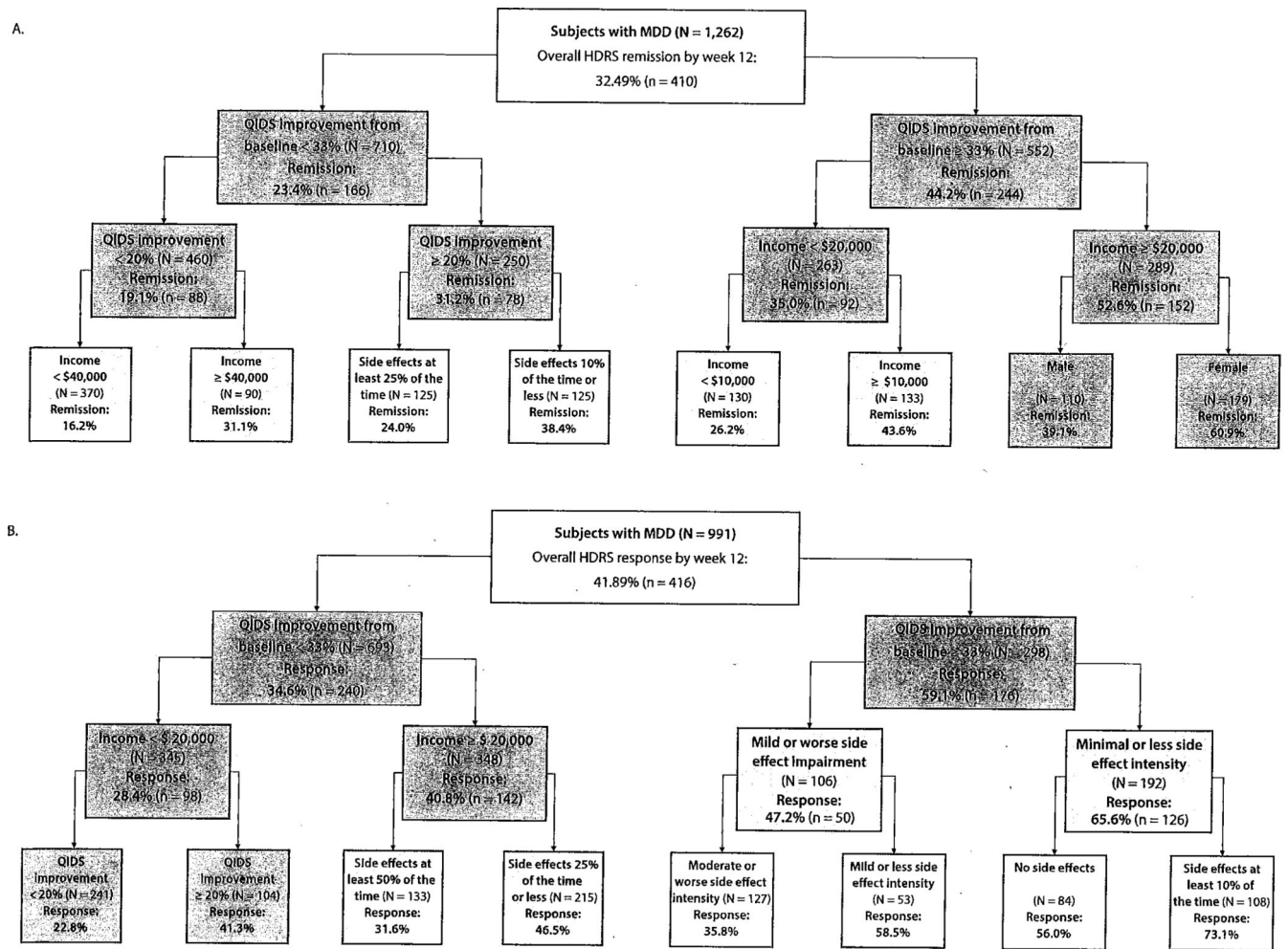operating characteristic.

Table 1

Clinical and Demographic Characteristics of Sample by Model

	Baseline				2 Weeks				4 Weeks			
Continuous Variables	Response (N=2,512)		Remission (N = 2,477)		Response (n=1,641)		Remission (n=1,857)		Response (n = 991)		Remission (n=1,262)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age, y	41.70	13.06	41.66	13.04	42.18	1.13	41.75	13.06	42.65	13.45	42.23	13.07
Age at onset of MDD, y	25.23	14.17	25.20	14.14	25.55	14.37	24.98	14.04	25.39	14.83	25.05	14.40
Length of MDD, y	16.44	13.29	16.44	13.30	16.60	13.52	16.74	13.36	17.23	13.73	17.18	13.51
Years of education	13.73	3.22	13.73	3.22	13.78	3.17	13.85	3.21	13.79	3.16	13.89	3.17
Categorical variables												
Income category[a]	1.49	1.13	1.51	1.13	1.49	1.14	1.49	1.13	1.46	1.14	1.47	1.13
Week 2												
Side effect intensity[b]					1.98	1.66	1.97	1.64	1.89	1.58	1.89	. 1.58
Side effect frequency[b]					2.02	1.90	2.01	1.88	1.96	1.87	1.94	1.83
Side effect burden[b]					1.37	1.48	1.33	1.44	1.26	1.35	1.23	1.31
Symptom improvement[c]					1.27	1.04	1.49	1.08	1.08	1.01	1.40	1.07
Week 4												
Side effect intensity[b]									1.84	1.64	1.83	1.62
Side effect frequency[b]									1.88	1.89	1.88	1.88
Side effect burden[b]									1.28	1.41	1.25	1.39
Symptom improvement[c]									1.34	1.06	1.69	1.07
Binary variables	n	%	n	%	n	%	n	%	n	%	n	%
Employed	1,563	63	1,563	63	990	61	1,139	62	581	60	758	61
Male gender	918	37	904	36	620	38	688	37	376	38	470	37
Hispanic	291	11	289	12	182	11	206	11	110	11	139	11
Race												
White	2,035	81	2002	81	1,360	83	1,543	83	817	82	1049	83
Black	416	16	414	17	249	15	277	15	152	15	187	15

Continuous Variables	Baseline				2 Weeks				4 Weeks			
	Response (N=2,512)		Remission (N = 2,477)		Response (n=1,641)		Remission (n=1,857)		Response (n = 991)		Remission (n=1,262)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Family history of MDD	1,393	56	1,368	56	906	56	1,046	57	546	55	710	57
Suicide attempts	423	17	418	17	278	17	323	17	168	17	222	18
Axis I comorbidity												
OCD	272	11	270	11	167	10	192	10	94	10	129	10
PTSD	420	17	416	17	276	17	318	17	173	18	227	18
Bulimia	307	12	305	12	210	13	238	13	132	13	185	15
Panic disorder	265	11	261	11	170	10	190	10	100	10	122	10
Agoraphobia	233	09	228	09	144	09	171	09	79	08	108	09
Social phobia	701	28	694	28	460	28	545	29	289	29	377	30
Alcohol dependence	254	10	251	10	172	11	190	10	115	12	134	11
Substance dependence	158	06	156	06	103	06	118	06	64	06	77	06
GAD	498	20	490	20	340	21	389	21	208	21	260	21
Somatization disorder	47	02	47	02	33	02	34	02	18	02	24	02
Hypochondriasis	89	04	88	04	61	04	66	04	32	03	38	03
Primary care	913	36	901	36	546	33	615	33	302	30	377	30
Marital status												
Married	888	35	871	35	592	36	656	35	356	36	443	35
Never married	716	29	710	29	454	28	530	29	270	27	351	28
Divorced	479	19	474	19	312	19	358	19	189	19	247	20
Widowed	76	03	74	03	56	03	61	03	38	04	44	03
Cohabitant	197	08	196	08	122	07	140	08	74	07	94	07
Separated	155	06	151	06	105	06	112	06	64	06	83	07
Academic degree												
Lower than high school	266	11	263	11	169	10	181	10	97	10	120	10
High school degree	1,515	60	1,494	60	986	60	1,116	60	602	61	755	60
Higher than high school	730	29	719	29	486	30	560	30	292	29	387	31
History of SSRI resistance	204	09	203	09	133	09	147	09	92	11	117	10
History of antidepressant resistance	272	11	268	11	183	12	200	11	128	13	159	13

Continuous Variables	Baseline				2 Weeks				4 Weeks			
	Response (N=2,512)		Remission (N = 2,477)		Response (n=1,641)		Remission (n=1,857)		Response (n = 991)		Remission (n=1,262)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Response												
QIDS (week 2)							248	13			138	11
QIDS (week 4)											232	18
HDRS (week 12)	1,339	53			783	48			416	42		
Remission												
QIDS (week 2)					32	02			21	02		
QIDS (week 4)									29	03		
HDRS (week 12)			1,023	41			709	38			410	32

[a] Categories: 0=< $10,000; 1 = $10,000–$20,000; 2 = $20,000–$40,000; and 3 = > $40,000.

[b] Side effects measured on a 7-point scale from 0 = none to 6 = intolerable.

[c] QIDS improvement measured in categories: −1 = decrease, 0 = 0%–10%, 1 = 10%–20%, 2 = 20%–33%, 3 = 33%–49%.

Abbreviations: GAD = generalized anxiety disorder, HDRS = Hamilton Depression Rating Scale, MDD = major depressive disorder, OCD = obsessive-compulsive disorder, PTSD = posttraumatic stress disorder, QIDS = Quick Inventory of Depressive Symptomatology, SSRI = selective serotonin reuptake inhibitor.