BMJ Open

What are the treatment remission, response and extent of improvement rates after up to four trials of antidepressant therapies in real-world depressed patients? A reanalysis of the STAR*D study’s patient-level data with fidelity to the original research protocol

H Edmund Pigott 1, Thomas Kim 2, Colin Xu 2, Irving Kirsch 3, Jay Amsterdam 4

ABSTRACT

Objective Reanalyse the patient-level data set of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study with fidelity to the original research protocol and related publications.

Design The study was open label and semirandomised examining the effectiveness of up to four optimised and increasingly aggressive, antidepressant therapies in depressed adults. Patients who failed to gain adequate relief from their level 1 trial on the SSRI citalopram could receive up to three additional treatment trials in levels 2–4.

Setting 41 North American psychiatry and primary care treatment centres.

Participants 4041 adults screened positive for major depressive disorder. In contrast to most clinical trials, STAR*D enrolled patients seeking care (vs recruited) and included patients with a wide range of common comorbid medical and psychiatric conditions to enhance the generalisability of findings to real-world clinical practice.

Interventions STAR*D evaluated the relative effectiveness of 13 antidepressant therapies in treatment levels 2–4 for depressed patients who failed to gain adequate benefit from their level 1 medication trial.

Main outcome measures According to the STAR*D protocol, the primary outcome was remission, defined as a score <8 on the blinded Hamilton Rating Scale for Depression (HRSD). Response was a secondary outcome defined as ≥50% reduction in HRSD scores. STAR*D’s protocol specifically excluded all non-blinded clinic-administered assessments from use as research outcome measures.

Results STAR*D investigators did not use the protocol-stipulated HRSD to report cumulative remission and response rates in their summary article and instead used a non-blinded clinic-administered assessment. This inflated their report of outcomes, as did their inclusion of 99 patients who scored as remitted on the HRSD at study outset as well as 125 who scored as remitted when initiating their next-level treatment. These should have been excluded from data analysis. In contrast to the STAR*D-reported 67%

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ We reanalysed the largest ever prospective antidepressant trial’s patient-level data set with fidelity to the original research protocol and related publications.

⇒ The reanalysis was conducted under the guidelines of the Restoring Invisible and Abandoned Trials initiative.

⇒ Treatment remission, response and extent of symptom improvement rates were calculated for 14 antidepressant therapies for those patients who met Sequenced Treatment Alternatives to Relieve Depression (STAR*D)’s inclusion in data analysis criteria as well as the overall cumulative remission rate after up to four trials of antidepressant therapies.

⇒ We calculated STAR*D’s remission rate using the protocol-stipulated Hamilton Rating Scale for Depression (HRSD) as well as combining the HRSD remissions with those from a non-stipulated measure of remission for patients missing an exit HRSD score. Combining STAR*D’s HRSD-defined remissions with those from the non-stipulated measure increased its cumulative remission rate from 35.0% to 41.3%.

⇒ Finally, we compared STAR*D’s outcomes to those found in a meta-analysis of 7030 patients enrolled in similar open-label antidepressant comparator trials, whereas the treatment remission and response rates in comparator trials averaged 48.4% and 65.2%, respectively, they were only 25.5% and 40.5% for STAR*D’s level 1 patients and worse in treatment levels 2–4. Similarly, comparator trials’ patients’ mean change on the HRSD was 14.8 points versus 8.4 points for STAR*D’s level 1 patients and worse for patients in treatment levels 2–4.
cumulative remission rate after up to four antidepressant treatment trials, the rate was 35.0% when using the protocol-stipulated HRSD and inclusion in data analysis criteria.

Conclusion: STAR*D’s cumulative remission rate was approximately half of that reported.

INTRODUCTION

At a cost of 35 million US dollars, the National Institute of Mental Health (NIMH) funded Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study is the largest and most expensive prospective antidepressant trial ever conducted with over 100 journal articles published by study investigators.1-7 In contrast to most clinical trials that enrol symptomatic volunteers (typically recruited through advertising), STAR*D enrolled 4041 patients who screened positive for major depressive disorder (MDD) while seeking routine medical or psychiatric care. STAR*D did not exclude patients with medical conditions and most comorbid psychiatric disorders, thereby increasing the generalisability of its findings to real-world clinical practice.

The STAR*D study provided up to four treatment trials per patient and was designed to give guidance in selecting the best next-level treatment option for the many patients who fail to gain sufficient relief from their first, and/or subsequent, antidepressant trial(s). To mimic clinical practice, STAR*D used an open-label research design with no control group during any phase of the study.

Our STAR*D reanalysis examines key methodological deviations from its research protocol and related publications, and these deviations impact on its investigators’ report of outcomes. In STAR*D’s Rationale and Research Design article, and repeated in the level 1–4 published study outcomes, STAR*D investigators stated, ‘the primary outcome is depressive symptom severity, measured by the 17-item Hamilton Rating Scale for Depression (HRSD)’ (Rush et al,1 p120). STAR*D’s prespecified primary outcome was remission, defined as scoring <8 on the HRSD, which was administered telephonically by Research Outcome Assessors (ROAs) blind to patients’ study status (treatment-level entry/exit/follow-up). Response was a secondary outcome defined as a ≥50% reduction in patients’ HRSD scores. Remission as defined by the HRSD (according to the protocol) was not presented in STAR*D’s summary article.2 Furthermore, despite its investigators’ numerous publications, neither change in HRSD depressive symptom severity nor HRSD response rates have been reported for STAR*D’s six primary studies4-6 and summary article.7 Instead, response rates and change in symptom severity were reported using the clinic-administered Quick Inventory of Depressive Symptomatology–Self Report (QIDS-SR), a measure developed by the STAR*D principal investigators.9 This occurred despite the fact that STAR*D’s research protocol specifically excluded all clinic-administered assessments, such as the QIDS-SR, from use as research outcome measures since they were not blinded and instead, used to guide patient care. The protocol states:

Recall that the research outcomes assessments are distinguished from assessments conducted at clinic visits. The latter are designed to collect information that guides clinicians in the implementation of the treatment protocol. Research outcomes assessments are NOT collected at the clinic visits. They are not collected by either clinicians or Clinical Research Coordinators (National Institute of Mental Health,10 p47,48; emphasis in the original).

In their summary article, STAR*D investigators used the QIDS-SR as the sole measure to report remission, response and extent of symptom improvement. The Abstract section of this article states that ‘the overall cumulative remission rate was 67%’ with no qualifiers to this claim (Rush et al,7 p1905). Besides making this claim based on an assessment the protocol specifically excluded from use as a research measure, it is not until the article’s Results section that careful readers learn this high level of treatment success did not occur. The STAR*D investigators’ claim was theoretical—an estimate based on the provisions of what would have happened if there were no study dropouts, and furthermore, ‘that those who exited the study would have had the same remission rates as those who stayed in the protocol’ (Rush AJ et al7 p1910).

As Pigott et al documented though, the investigators’ assumptions are not true in the real world since more patients dropped out than remitted in each STAR*D treatment level,11 and furthermore, it has been found in placebo-controlled trials that patients who drop out are more likely to have had adverse treatment side effects and/or emergent suicidality.12

Unfortunately, the STAR*D investigators’ claim of a 67% cumulative remission rate has become accepted clinical wisdom, and the provisions on which it is based are commonly not referenced when portraying STAR*D’s findings. For example, in 2009, NIMH’s director Dr Thomas Insel claimed that STAR*D found ‘at the end of 12 months, with up to four treatment steps, roughly 70% of participants were in remission’ (Insel and Wang13 p1406). Similarly in 2013, an editorial in the American Journal of Psychiatry (AJP) claimed that STAR*D found ‘after four optimised, well-delivered treatments, approximately 70% of patients achieve remission’ (Greden,14 p580). More recently (2022), a New York Times’ article claimed that half of STAR*D’s participants ‘had significantly improved after using either the first or second medication, and nearly 70% of people had become symptom free by the fourth antidepressant’.15 These are not factual statements of STAR*D’s findings.

The first author has made published criticisms alleging protocol violations that appear to inflate STAR*D’s findings and called for the reanalysis of the data set by independent investigators.16 In 2018, the first and fourth authors collaborated with researchers from the University of Connecticut to reanalyse STAR*D’s level 1 data obtained from NIMH.17 This reanalysis found substantial inflation of STAR*D’s reported remission and response
rates. Furthermore, the reanalysis found that the extent of HRSD improvement in STAR*D’s level 1 trial was approximately half that of open-label antidepressant comparator trials.

Our published criticisms of STAR*D investigators’ report of outcomes are as follows.

► While STAR*D investigators used the HRSD to report remission rates in their levels 1–4 articles, the QIDS-SR was used as the sole measure to report remission, response and extent of improvement rates in their summary article without disclosing that the protocol specifically excluded all non-blinded/clinic-administered assessments such as the QIDS-SR from use as outcome measures. The primary outcome measure, the HRSD, should have been used to report the summary article’s outcomes.

► Using data from the 931 patients deemed ineligible for analysis in STAR*D’s level 1 article because these patients lacked a baseline ROA-administered HRSD score ≥14, in STAR*D’s levels 2–4 and summary articles without clear disclosure. This included 99 patients who scored <8 on their baseline HRSD—indicating these patients met STAR*D’s remission criterion at study outset and should not have been included in their report of outcomes.

► Excluding from analysis, 370 patients who dropped out after starting on citalopram in their first clinic visit without taking the exit HRSD despite STAR*D investigators stating, ‘our primary analyses classified patients with missing exit HRSD scores as nonremitters a priori’ (Trivedi et al., p34). These 370 early dropout patients should have been counted as non-remitters as prespecified.

► Including in their analyses, 125 patients who scored as remitted at entry into their next-level treatment. This occurred despite STAR*D investigators prespecifying that ‘patients who begin a level with HRSD<8 will be excluded from analyses’ (Rush et al., p130).

This reanalysis article uses the patient-level data set obtained from NIMH to replicate the STAR*D summary article, which used descriptive statistics to present the remission, response and extent of symptomatic improvement for 14 antidepressant therapies based on the QIDS-SR. We perform the same descriptive analyses with the key differences compared with those presented in STAR*D’s summary article being: (1) ours is based on the protocol-specified HRSD and only uses the QIDS-SR for those patients missing their exit HRSD and (2) we only included patients who met the inclusion for data analysis criteria stipulated in the research protocol and related publications. Future efforts will use inferential statistics to reanalyse STAR*D’s levels 2–4 semirandomised comparator trials, including the extent of emergent suicidal ideation and 12-month follow-up outcomes tied to each compared treatment.

METHOD
Restoring Invisible and Abandoned Trials initiative
The Restoring Invisible and Abandoned Trials (RIAT) initiative started in 2013 calling on funders and investigators of abandoned (unpublished) or misreported studies to publish undisclosed outcomes or correct misleading publications. If investigators failed to correct a study identified as misreported, independent investigators were encouraged to correct the record by reanalysing the study’s patient-level data set consistent with the research protocol and analytic plan.

On 6 March 2019, the RIAT investigators published our response to a ‘Call to Action’ statement in the British Medical journal, in which we stated our intention to reanalyse the STAR*D data set. We then notified STAR*D’s principal investigators of our intention and requested they inform us whether they would undertake a reanalysis of the data set adhering to the research protocol. On 22 March 2019, STAR*D investigators acknowledged our email notification, indicated that the STAR*D data were in the public domain, and stated they had no interest in undertaking a reanalysis.

In July 2019, we received an STAR*D Data Use Certificate, issued by the NIMH Data Access Committee, and gained access to the STAR*D levels 1–4 and follow-up patient-level data set consisting of 26 text files and limited supporting study documentation. In September 2019, we obtained funding from the RIAT Support Center to reanalyse STAR*D.

Patients
STAR*D patients were 18–75 years of age, seeking care at 18 primary and 23 psychiatric care clinics. Clinical research coordinators (CRCs) screened 4790 patients for MDD. This screening included the CRCs’ administering the HRSD, on which 4041 patients scored ≥14, met the other inclusion criteria, and enrolled into the study. CRCs also gathered patients’ psychiatric history, demographic information and administered both the Cumulative Illness Rating Scale and the Psychiatric Diagnostic Screening Questionnaire to determine the extent of comorbid medical and psychiatric disorders.

Levels/steps of acute treatment
STAR*D investigators sought to provide the highest quality of care to maximise the number of remissions while minimising dropouts (see online supplemental table 1). Online supplemental table 2 describes the antidepressant therapies available in treatment levels 1–4 while steps refer to the numeric order of treatments. As seen in figure 1, treatment steps 1 and 2 correspond to levels 1 and 2 treatments. Similarly, for most patients, their levels 3 and 4 treatments correspond to treatment steps 3 and 4. For level/step 2 patients though who failed to respond adequately to cognitive therapy alone or combined with citalopram and chose to continue in the study, their third treatment step was designated level 2A and they were randomised to one of two level 2 switch
medications. For these patients, their level 2A treatment was their third treatment step. For level 2A patients who did not adequately benefit from this medication trial and chose to continue in the study, they entered a fourth treatment step consisting of level 3 treatments. All patients were administered the SSRI citalopram for their level 1 treatment. Each treatment level consisted of 12 weeks of antidepressant therapy, with an additional 2 weeks for patients deemed close to remission. Treatment was administered using a system of measurement-based care that assessed symptoms and side effects at each clinic visit. STAR*D investigators state, ‘To enhance the quality and consistency of care, physicians used the clinical decision support system that relied on the measurement of symptoms (QIDS-C and QIDS-SR), side-effects, medication adherence, and clinical judgment based on patient progress’ (Trivedi et al,1 p30). This system was used to guide medication management of a fully adequate dose for a sufficient time to ‘ensure that the likelihood of achieving remission was maximized and that those who did not reach remission were truly resistant to the medication’ (Trivedi et al,1 p30).

For those patients who failed to gain an adequate response from citalopram, STAR*D allowed them to select acceptable treatment options for randomisation in levels 2 to 4 ‘to empower patients, strengthen the therapeutic alliance, optimize treatment adherence, and improve outcome’ (Fava et al,20 p483). The treatment options available for randomisation involved either switching to a new treatment or augmenting the patient’s current treatment. Treatment levels 2–4 evaluated the relative effectiveness of 11 pharmacologically distinct drug/drug combination treatments. Cognitive therapy was also available as either a switch or citalopram augmentation option in level 2.
STAR*D follow-up phase

In each treatment trial for levels 1–4, patients who scored <6 on their last QIDS–Clinician version (QIDS–C) were considered clinician-rated remissions and encouraged to enter the 12-month follow-up phase. During follow-up, patients continued their ‘previously effective acute treatment medication(s) at the doses used in acute treatment but that any psychotherapy, medication, or medication dose change could be used’ (Rush et al., p1908). Based on prior research, a QIDS score of <6 was estimated by STAR*D investigators to correspond to a score of <8 on the HRSD. STAR*D’s prespecified primary outcome measure for classifying patients as remitted. Clinicians strongly encouraged patients who did not obtain a QIDS-defined remission to enter the next-level treatment. Patients who failed to attain a QIDS-defined remission but did have a ≥50% reduction on the QIDS–C did not want to be randomised to a next-level treatment were also encouraged to enter follow-up.

Research design of the STAR*D study

STAR*D investigators developed a new research design for the study termed ‘equipoise-stratified’ to evaluate the relative efficacy of 13 antidepressant therapies in levels 2–4 for depressed patients who failed to gain adequate benefit from their level 1 medication trial. In level 1, all patients received citalopram as their first treatment. In level 2, patients were informed regarding seven treatment options to choose from: four switch options in which citalopram was stopped and the new treatment initiated and three augmentation options in which citalopram was combined with a second antidepressant treatment. In level 3, patients were informed regarding four treatment options to choose from: two switch options and two augmentation options. Level 4 involved randomisation to one of two medication/medication combination switch options.

Analytic plan of the RIAT reanalysis

We reanalysed the STAR*D patient-level data set with fidelity to the original research protocol whenever possible. Where the protocol was silent, we used other STAR*D publications to guide our analysis. This occurred four times. First, the protocol is silent regarding patients who entered the study without a baseline ROA-administered HRSD score of ≥14. In their level 1 article, STAR*D investigators deemed the 931 such patients who lacked this marker of depression severity ineligible for inclusion in data analysis. We do the same and extend this exclusion for such patients who continued on to levels 2–4 because their extent of depression severity at study onset is not known. Second, the protocol is silent on what to do with patients who met the remission criteria on the HRSD at entry into their next-level treatment. In STAR*D’s Rationale and Research Design article though, its investigators prespecify that ‘patients who begin a level with HRSD <8 will be excluded from analyses’ (Rush et al., p130). We, therefore, excluded 125 such patients from our analyses of treatment levels 2–4. Third, the protocol is silent on how to analyse patients who exit a treatment without taking the HRSD. STAR*D investigators state in their level 1 article, ‘our primary analyses classified patients with missing exit HRSD scores as nonremitters a priori’ (Trivedi et al., p34) and repeat similar statements in their level 2–4 articles. Therefore, we do likewise.

Finally, STAR*D had many patients with missing exit HRSD scores. In their level 2–4 articles, STAR*D investigators used a correspondence table to map the final QIDS–SR score to the HRSD for patients missing their exit HRSD score to assess the impact of their approach to counting such patients as ‘nonremitters a priori’. For patients with missing exit HRSD scores, we therefore mapped their last QIDS–SR score to the HRSD and used it to calculate the mean HRSD exit, mean change and combined HRSD and QIDS–SR response rates for all treatments. We also calculated STAR*D’s remission rate both as prespecified based on an exit HRSD score of <8 as well as a final QIDS–SR score of <6 for those patients missing an exit HRSD score.

All preprocessing and analyses were performed in R. Authors 2 and 3 identified patients by their subject key and used this variable to match information across data sets. Data on patients’ treatment pathways, and when patients transitioned from one level to the next, were taken from the Integrated Voice Response Alert data set completed by CRCs, and verified against the data on patient-level exits. Authors 2 and 3 then compared the number of patients identified for all level 1–4 treatments to that reported in the STAR*D summary article’s patient flowchart and the number of patients matched.

Next, authors 2 and 3 applied STAR*D’s level 1 inclusion criteria for data analysis criterion to patients in treatment levels 2–4 as well as excluded from analysis the 125 patients who scored <8 on the HRSD at entry into their next-level treatment. We counted these 125 patients as remitted in the prior treatment level but excluded them from the analyses of subsequent treatments. Online supplemental table 3 presents the number of level 2–4 patients excluded from our reanalysis, and the reasons for their exclusion. Online supplemental table 4 identifies the number of patients with missing entry and/or exit HRSD scores for all level 1–4 treatments. As seen in online supplemental table 4, 1390 patients were missing their exit HRSD score across all treatments.

We then compared STAR*D’s outcomes to those found in a meta-analysis of 7030 patients enrolled in antidepressant comparator trials. Similar to STAR*D, comparator trials typically are conducted open-label without a control group and, therefore, are the appropriate comparison data for STAR*D’s outcomes. Continuous HRSD improvement means were provided by the first author of the meta-analysis.

Finally, we compared the STAR*D protocol’s step-by-step predictions of patient drop out and the number of patients who would have a satisfactory treatment response and enter follow-up care to what actually occurred.
While the purpose of these predictions’ was to estimate the number of continuing patients available for randomisation in treatment levels 2–4, at the meta-level, these predictions are an important hypothesis STAR*D tested by assessing how well its investigators could predict the aggregate step-by-step successful treatment outcomes from their treat-to-remission model of care.

Patient and public involvement
Neither patients nor the public were involved in the design, conduct, reporting or dissemination plans of our research.

RESULTS
Figure 1 presents the overall flow of patients enrolled in the various protocol-defined treatment levels and places them in groups defined by the number of treatment steps. Of the 4041 patients enrolled into STAR*D, 3110 met the eligibility for data analysis criterion of having an ROA-administered HRSD score ≥14 at study outset. Figure 1 also identifies the number of patients who exited the study following each treatment step, the number who entered follow-up after each treatment step and the number who were randomly assigned to a next-level treatment.

Online supplemental table 5 describes the demographic and clinical features of the patients who entered treatment in steps 1–4 based on their level 1 baseline presentation when enrolling into the study. Summary statistics are presented as means and SDs for continuous variables and percentages for discrete variables. Note that 55.7% of STAR*D patients had two or more comorbid axis 1 disorders when first enrolled based on the Psychiatric Diagnostic Screening Questionnaire and averaged 2.5 comorbid medical conditions based on the Cumulative Illness Rating Scale. Furthermore, the average length of patients’ current MDD episode was 25.9 months. In a post hoc analysis, STAR*D investigators found that 77.8% of its enrolled patients would have been excluded from most antidepressant trials due to having two or more concurrent medical conditions, more than one comorbid psychiatric disorder and/or a current depressive episode lasting >2 years.26

Table 1 presents the mean HRSD entry, exit and change scores for patients by the specific treatment they received in steps 1–4 as well as the HRSD remission and response rates. Table 1 also provides the HRSD cumulative remission rate after up to four trials on antidepressant therapies as well as the combined HRSD plus QIDS-SR remission and response rates for the 1330 patients missing an exit HRSD score.

Table 2 presents patients’ aggregate HRSD status in terms of remission, response and extent of mean symptomatic change at entry and exit for each treatment step as well as study dropout. In step 1, 25.5% of patients remitted. Steps 2–4 show a continuous decrease in remission rates from step 2’s 21.3% to step 3’s 13.2% and step 4’s 10.4% with increasing rates of study dropout from step 1’s 34.5% to step 3’s 46.2%.

Online supplemental figure 1 and online supplemental figure 2 compare the HRSD remission, response and extent of symptom improvement rates for STAR*D patients in steps 1–4 to that found in a meta-analysis of 7030 patients enrolled in non-blinded antidepressant comparator trials.25 In step 1, these measures of improvement among STAR*D’s patients were at least one-third less than that found in comparator trials, and improvement was worse in each subsequent treatment step.

Figure 2 compares the STAR*D protocol’s predictions of patient dropout and the number of patients who would have a satisfactory treatment response and enter follow-up to what occurred. Cumulatively, STAR*D’s investigators predicted that 73.8% of patients would have a successful treatment response and enter follow-up, whereas in fact only 45.6% achieved this measure of treatment success. Furthermore, whereas STAR*D investigators predicted that over the course of up to four antidepressant therapies, 20.7% of patients would dropout, in fact, 53.7% dropped out. On this measure of treatment failure, STAR*D’s dropout rate was 2.6 times greater than predicted.

Figure 3 presents the step-by-step cumulative remission rate in three ways. First, the ‘theoretical’ rate propagated by STAR*D investigators based on the provisos of what would have happened if there were no study dropouts and that those who did exit had the same QIDS-SR remission rates as those who stayed.7 Next, the combined HRSD plus QIDS-SR remission rate based on either an exit HRSD score of <8, OR a last clinic visit QIDS-SR score of <6 for the 1330 patients missing an exit HRSD. Finally, the RIAT reanalysis rate when using the protocol-specified exit HRSD score of <8 as the sole measure of remission for the 3110 patients who met STAR*D’s inclusion in data analysis criteria. The cumulative remission rate after up to four antidepressant therapies using the HRSD was 35.0% vs 41.3% when combined with the QIDS-SR, both of which are substantially less than the 67% cumulative remission rate claimed in the summary article’s Abstract.

DISCUSSION
Principal findings and comparison with original STAR*D publication
STAR*D’s results highlight the discrepancy in likely outcomes between typical antidepressant clinical trials with their exclusion criteria and the real-world patients for whom these medications are commonly prescribed. Our RIAT reanalysis found poorer outcomes after up to four optimised, and increasingly aggressive, antidepressant therapies than reported in STAR*D’s summary article published in AJP.7 In contrast to the 67% cumulative remission rate reported in AJP, the actual rate was 35.0% when using the protocol-specified HRSD and increased to 41.3% when combined with a final clinic-visit QIDS-SR score of <6 for patients’ missing exit HRSD scores in treatment steps 1–4. The 41.3% cumulative remission rate should be viewed as
### Table 1  Outcomes across all treatments

<table>
<thead>
<tr>
<th>Treatment step</th>
<th>HRSD Score</th>
<th>*Mean change (95% CI) (SD)</th>
<th>HRSD remission rate # (%)</th>
<th>*Combined HRSD and QIDS-SR remission rate # (%)</th>
<th>*Combined HRSD and QIDS-SR response rate # (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 (N=3110)</td>
<td>Entry mean (SD)</td>
<td>Exit mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2 (N=1134)</td>
<td>21.87 (5.21)</td>
<td>13.49 (8.42)</td>
<td>8.38 (8.10, 8.67) (8.11)</td>
<td>794 (25.5%)</td>
<td>938 (30.2%)</td>
</tr>
<tr>
<td>Switch strategy (N=620)</td>
<td>18.76 (6.24)</td>
<td>13.97 (8.09)</td>
<td>4.79 (4.37, 5.21) (7.23)</td>
<td>241 (21.3%)</td>
<td>283 (25.0%)</td>
</tr>
<tr>
<td>Bupropion (N=190)</td>
<td>20.11 (6.08)</td>
<td>14.70 (8.01)</td>
<td>5.16 (4.59, 5.73) (7.22)</td>
<td>113 (18.2%)</td>
<td>134 (21.6%)</td>
</tr>
<tr>
<td>Sertraline (N=198)</td>
<td>19.95 (5.98)</td>
<td>14.92 (8.02)</td>
<td>5.03 (4.04, 6.01) (7.10)</td>
<td>32 (16.2%)</td>
<td>36 (18.2%)</td>
</tr>
<tr>
<td>Venlafaxine (N=192)</td>
<td>19.89 (6.19)</td>
<td>14.31 (8.12)</td>
<td>5.58 (4.53, 6.63) (7.45)</td>
<td>37 (19.3%)</td>
<td>44 (22.9%)</td>
</tr>
<tr>
<td>Cognitive therapy (N=40)</td>
<td>18.01 (4.96)</td>
<td>12.44 (7.93)</td>
<td>5.58 (2.87, 8.28) (7.83)</td>
<td>13 (32.5%)</td>
<td>17 (42.5%)</td>
</tr>
<tr>
<td>Augmentation strategy (N=514)</td>
<td>17.44 (6.18)</td>
<td>13.10 (8.10)</td>
<td>4.34 (3.72, 4.97) (7.23)</td>
<td>128 (24.9%)</td>
<td>149 (29.0%)</td>
</tr>
<tr>
<td>Bupropion (N=216)</td>
<td>16.88 (6.11)</td>
<td>12.52 (7.83)</td>
<td>4.36 (3.38, 5.33) (7.30)</td>
<td>54 (25.0%)</td>
<td>64 (29.6%)</td>
</tr>
<tr>
<td>Buspirone (N=225)</td>
<td>17.80 (6.50)</td>
<td>13.36 (8.40)</td>
<td>4.43 (3.52, 5.35) (7.02)</td>
<td>58 (25.8%)</td>
<td>68 (30.2%)</td>
</tr>
<tr>
<td>Cognitive therapy (N=73)</td>
<td>17.99 (5.24)</td>
<td>13.98 (7.98)</td>
<td>4.01 (2.25, 5.78) (7.69)</td>
<td>16 (21.9%)</td>
<td>17 (23.3%)</td>
</tr>
<tr>
<td>Step 3 (N=325)</td>
<td>19.59 (6.09)</td>
<td>16.38 (7.77)</td>
<td>3.21 (2.48, 3.94) (6.70)</td>
<td>43 (13.2%)</td>
<td>50 (15.4%)</td>
</tr>
<tr>
<td>Level 2A (N=28)</td>
<td>20.89 (5.44)</td>
<td>16.96 (6.48)</td>
<td>3.93 (1.81, 6.04) (5.71)</td>
<td>3 (10.7%)</td>
<td>3 (10.7%)</td>
</tr>
<tr>
<td>Bupropion (N=12)</td>
<td>19.92 (3.85)</td>
<td>17.58 (7.35)</td>
<td>2.33 (−0.81, 5.48) (5.55)</td>
<td>2 (16.7%)</td>
<td>2 (16.7%)</td>
</tr>
<tr>
<td>Venlafaxine (N=16)</td>
<td>21.62 (6.41)</td>
<td>16.50 (5.96)</td>
<td>5.12 (2.33, 7.92) (5.70)</td>
<td>1 (6.2%)</td>
<td>1 (6.2%)</td>
</tr>
<tr>
<td>Level 3 (N=297)</td>
<td>19.46 (6.14)</td>
<td>16.32 (7.88)</td>
<td>3.14 (2.37, 3.92) (6.79)</td>
<td>40 (13.5%)</td>
<td>47 (15.8%)</td>
</tr>
<tr>
<td>Switch strategy (N=186)</td>
<td>20.01 (6.24)</td>
<td>17.01 (7.91)</td>
<td>2.99 (2.00, 3.99) (6.94)</td>
<td>23 (12.4%)</td>
<td>25 (13.4%)</td>
</tr>
<tr>
<td>Nortriptyline (N=92)</td>
<td>19.67 (5.27)</td>
<td>16.99 (8.35)</td>
<td>2.67 (1.10, 4.24) (7.68)</td>
<td>15 (16.3%)</td>
<td>15 (16.3%)</td>
</tr>
<tr>
<td>Mirtazapine (N=94)</td>
<td>20.34 (7.08)</td>
<td>17.03 (7.49)</td>
<td>3.30 (2.06, 4.55) (6.15)</td>
<td>8 (8.5%)</td>
<td>10 (10.6%)</td>
</tr>
</tbody>
</table>

Continued
the ‘best case scenario’ since it added an additional 195 QIDS-defined remissions (a remission measure not specified in the protocol) from the 1330 patients with missing exit HRSD scores. As there was neither a placebo nor waitlist control group during any phase of the STAR*D study, it is impossible to know to what extent the observed results were due to the pharmacologic effects of the prescribed medications, placebo effects and/or the passage of time.

Our reanalysis did not assess the durability of treatment effects during the 12-month follow-up phase. In their summary article though, STAR*D investigators reported an overall relapse rate of 46.1% for the 1729 patients who had at least one assessment (of up to 12 scheduled) during follow-up using a telephonic-administered version of the QIDS, whereas Pigott et al found a far lower sustained

### Table 1  
Continued

<table>
<thead>
<tr>
<th>Treatment step</th>
<th>HRSD Score</th>
<th>*Mean change (95% CI)</th>
<th>HRSD remission rate # (%)</th>
<th>*Combined HRSD and QIDS-SR remission rate # (%)</th>
<th>*Combined HRSD and QIDS-SR response rate # (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmentation strategy (N=111)</td>
<td>18.55 (5.89)</td>
<td>15.16 (7.74)</td>
<td>3.40 (2.18, 4.62)</td>
<td>17 (15.3%)</td>
<td>22 (19.8%)</td>
</tr>
<tr>
<td>Lithium (N=58)</td>
<td>18.69 (6.47)</td>
<td>15.91 (7.29)</td>
<td>2.78 (1.12, 4.15)</td>
<td>7 (12.1%)</td>
<td>9 (15.5%)</td>
</tr>
<tr>
<td>T3 (N=53)</td>
<td>18.41 (5.25)</td>
<td>14.34 (8.19)</td>
<td>4.07 (1.99, 6.14)</td>
<td>10 (18.9%)</td>
<td>13 (24.5%)</td>
</tr>
<tr>
<td>Step 4 (N=106)</td>
<td>20.65 (5.54)</td>
<td>16.49 (7.47)</td>
<td>4.16 (2.80, 5.52)</td>
<td>11 (10.4%)</td>
<td>13 (12.3%)</td>
</tr>
<tr>
<td>Level 3 (N=16)</td>
<td>20.62 (4.01)</td>
<td>17.62 (6.87)</td>
<td>3.00 (−0.45, 6.45)</td>
<td>2 (12.5%)</td>
<td>2 (12.5%)</td>
</tr>
<tr>
<td>Tranylcypromine (N=43)</td>
<td>21.02 (6.57)</td>
<td>16.45 (7.89)</td>
<td>4.57 (2.22, 6.92)</td>
<td>3 (7.0%)</td>
<td>5 (11.6%)</td>
</tr>
<tr>
<td>Venlafaxine XR/mirtazapine (N=47)</td>
<td>20.32 (5.02)</td>
<td>16.14 (7.38)</td>
<td>4.18 (2.30, 6.07)</td>
<td>6 (12.8%)</td>
<td>6 (12.8%)</td>
</tr>
</tbody>
</table>

Cumulative remission rate after up to four treatment steps: 1089 (35.0%) and 1284 (41.3%).

*For patients with missing exit HRSD scores, their last QIDS-SR score is mapped to the HRSD and used to calculate HRSD exit mean, mean change, combined HRSD and QIDS-SR remission rate and combined HRSD and QIDS-SR response rate.

HRSD, Hamilton Rating Scale for Depression; QIDS-SR, Quick Inventory of Depressive Symptomatology–Self Report.

### Table 2  
Outcomes by treatment step

<table>
<thead>
<tr>
<th>Step 1 (N=3110)</th>
<th>Step 2 (N=1134)</th>
<th>Step 3 (N=325)</th>
<th>Step 4 (N=106)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td><strong>SD</strong></td>
<td><strong>Mean</strong></td>
<td><strong>SD</strong></td>
</tr>
<tr>
<td>HRSD score at entry into step</td>
<td>21.87</td>
<td>5.21</td>
<td>18.76</td>
</tr>
<tr>
<td>HRSD score at exit from step*</td>
<td>13.49</td>
<td>8.42</td>
<td>13.97</td>
</tr>
<tr>
<td>HRSD mean change*</td>
<td>8.38</td>
<td>8.11</td>
<td>4.79</td>
</tr>
<tr>
<td>Remission at each step exit</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Response at each step exit*</td>
<td>1261</td>
<td>40.5%</td>
<td>329</td>
</tr>
<tr>
<td>Entered follow-up</td>
<td>902</td>
<td>29.0%</td>
<td>406</td>
</tr>
<tr>
<td>Study exit/dropout</td>
<td>1074</td>
<td>34.5%</td>
<td>403</td>
</tr>
</tbody>
</table>

*For patients with missing exit HRSD scores, their last QIDS-SR score is mapped to the HRSD and used to calculate HRSD exit mean, mean change and combined HRSD and QIDS-SR response rate.

HRSD, Hamilton Rating Scale for Depression; QIDS-SR, Quick Inventory of Depressive Symptomatology–Self Report.
recovery rate when incorporating patient dropout in the analysis.12

Comparison with other studies
Our reanalysis found that in step 1, STAR*D’s remission, response and extent of improvement rates were substantially less than those reported in other open-label antidepressant comparator trials and then grew progressively worse in steps 2–4.25 Such studies typically exclude depressed patients with the range and number of comorbid medical and/or psychiatric disorders that were included in STAR*D.

STAR*D’s step 1 remission rate was 25.5% followed by a progressive decline in remission rates for those patients receiving subsequent, and increasingly aggressive treatments, such that by step 4, it was only 10.4%. This decline in antidepressants’ effectiveness essentially mirrors the findings from randomised and naturalistic, prospective studies reporting a 20%–30% loss of effectiveness with each increase in the number of prior antidepressant trials.27–32 Furthermore, several recent analyses suggest that the sequential application of antidepressant

---

**Figure 2** Comparison of STAR*D protocol predictions to what occurred. RIAT, Restoring Invisible and Abandoned Trial; STAR*D, Sequenced Treatment Alternatives to Relieve Depression.

**Figure 3** STAR*D’s step-by-step cumulative remission rate presented three ways. The step-by-step theoretical remission rates were obtained from the STAR*D summary article where it states: ‘The theoretical cumulative remission rate is 67% (37+19+6+5).’ [Rush AJ et al, p1910]. The HRSD+QIDS-SR cumulative remission rate was taken from table 1. It combines the 1089 patients with an exit HRSD score of<8 with the 195 patients who were missing an exit HRSD score but had a final clinic-visit QIDS-SR score of<6. The RIAT Reanalysis cumulative remission rate is based on an exit HRSD score of<8 as the sole measure of remission for the 3110 patients who met STAR*D’s inclusion for data analysis criteria. HRSD, Hamilton Rating Scale for Depression; QIDS-SR, Quick Inventory of Depressive Symptomatology–Self Report; RIAT, Restoring Invisible and Abandoned Trial; STAR*D, Sequenced Treatment Alternatives to Relieve Depression.
medications for non-remitting depression may in fact foster treatment resistance for many patients.

Regarding the protocol’s predictions of treatment success and patient dropout, it states:

We arrived at these estimates using three experienced practitioners who independently made estimates that were surprisingly close to each other. Then, via tele-conferencing, the final estimates were made. The underlying assumptions of these estimates come largely by inferences from results of published RCTs (National Institute of Mental Health, p31; emphasis added).

STAR*D’s actual measures of treatment success and failure were significantly worse than predicted. As Barbui et al noted, antidepressant study dropout rates provide a ‘hard measure of treatment effectiveness and acceptability’ (Barbui et al., p296) and STAR*D’s dropout rate was 2.6 times greater than predicted. This discrepancy further highlights the relative ineffectiveness of antidepressants in treating real-world depressed patients, compared with those reported in conventional studies.

CONCLUSION

Bias in the clinical literature is commonly associated with industry-funded RCTs, not publicly funded ones. Our RIAT reanalysis though documents scientific errors in this NIMH-funded study. These errors inflated STAR*D investigators’ report of positive outcomes.

The STAR*D summary article’s claim of a 67% cumulative remission rate was published in 2006. If STAR*D’s outcomes had been reported as prespecified, its model of care would likely have faced much stronger criticism 17 years ago and fueled a more vigorous search for evidence-based treatment alternatives.

REFERENCES


15 Smith DG. Antidepressants don't work the way many people think. New York Times 8 November 2022.
16 Pigott HE. The STAR*D trial: it's time to reexamine the clinical beliefs which guide the treatment of major depression. Can J Psychiatry 2015;60:9–13.
32 Amsterdam JD, Lorenzo-Luaces L, DeRubeis RJ. Step-wise loss of antidepressant effectiveness after repeated antidepressant trials in bipolar II depression. Bipolar Disord 2016;18:563–70.
Comparator trial meta-analysis (N=7,030)
Step 1 (N=3,110)
Step 2 (N=1,134)
Step 3 (N=325)
Step 4 (N=106)

Percent (%)

- Remission rate
- Response rate

Remission rate: 48.4% (Comparator trial), 25.5% (Step 1), 21.3% (Step 2), 13.2% (Step 3), 10.4% (Step 4)
Response rate: 65.2% (Comparator trial), 40.5% (Step 1), 29.0% (Step 2), 19.4% (Step 3), 20.8% (Step 4)
Step 1 (N=3,110)
Step 2 (N=1,134)
Step 3 (N=325)
Step 4 (N=106)

Points of change (HRSD scale)
Comparator trial meta-analysis (N=7,030)
### Supplementary Table 1:
**Highest Quality of Acute and Continuing-Care to Maximize Remissions While Minimizing Relapse and Dropouts**

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Explanation</th>
</tr>
</thead>
</table>
| Optimized Sustained Study Participation to Minimize Dropouts | • Promoted patients’ study affiliation via STAR*D-branded brochures, bimonthly newsletters, and an informational video emphasizing STAR*D’s public health significance and the critical role played by patients;  
• Educated patients and families about depression and its treatment using a multi-step educational package. This included teaching the “mechanism of action” for patients’ current antidepressant and educating patients that “depression is a disease, like diabetes or high blood pressure” and “can be treated as effectively as other illnesses,” etc.;  
• Used a letter reminder system to alert patients before appointments in those clinics without such systems who had a >15% rate of missed appointments;  
• Ensured timely follow-up and rescheduling of missed appointments by calling patients on the day of the missed appointment, and again within 24 hours, if there was no response. Patient’s physician sent letter within 48 hours if contact was not established;  
• Used a letter reminder system for all research outcome assessment calls during acute and continuing-care;  
• In every clinic visit, the Clinical Research Coordinator (CRC) discussed the research outcomes phone calls with the patient to ensure that the calls were completed on schedule and worked to resolve any problematic issues regarding said calls [Clinical Procedures Manual, page 75];  
• Paid patients $25.00 for participating in each telephonic research outcomes assessment;  
• Permitted patients to re-enter acute and/or continuing-care within four weeks after having dropped out [Clinical Procedures Manual, page 80];  
• Recommended one-year of continuing-care for all patients who achieved a satisfactory clinical response with the essential goal of preventing relapse [Clinical Procedures Manual, page 15] and  
• Permitted continuing-care patients to remain in the study if they moved from the area [Clinical Procedures Manual, page 81]. |
| Acute-Care Visits | Physicians met with patients on entry into each new step to initiate drug treatment with follow-up visits scheduled on weeks 2, 4, 6, 9, 12, with an optional week 14 visit. |
Measurement-Based Care

Conducted structured evaluations of symptoms and side-effects at each visit and included a centralized treatment monitoring and physician feedback system to ensure consistent implementation of optimal care across research sites.

Aggressive Medication Dosing

Provided aggressive medication dosing with a fully adequate dose for a sufficient duration to “ensure that the likelihood of achieving remission was maximized and that those who did not reach remission were truly resistant to the medication”, 1, p.30

Liberal Prescribing of Non-Study Medications

Physicians had great leeway in prescribing non-study medications to treat comorbid symptoms resulting in:

- 17.2% taking Trazodone for sleep;
- 11.9% taking an anti-anxiety medication;
- 16.7% taking either a sedative or hypnotic medication; and
- An undisclosed percent taking medications to address side-effects. 2, table 2

Continuing-Care Visits

Patients saw their physician every 2 months and continued taking their treatment medication(s) at the same doses but their physicians were allowed to make any psychotherapy, medication, and/or medication dose changes to maximize the likelihood of maintaining patients’ remission status. 7, p. 1908 Additional continuing-care visits were scheduled when patients began to experience a return of depressive symptoms and/or intolerable side-effects [Clinical Procedures Manual. page 78].

Clinical Research Coordinator (CRC)

Each site had a CRC who: 1, p. 30

- Saw patients before each visit administering multiple measures to them including the QIDS-SR during each acute-care visit;
- Assisted physicians in protocol implementation; and
- Provided patients support and encouragement in protocol implementation.

Treatment Designed to Enhance Subject Retention

Treatment was designed to minimize drop-outs and/or non-compliance including:

- Open label prescribing during acute and continuing-care with no placebo control condition during any study phase;
- Patients chose their acceptable treatment assignments for steps two and three to eliminate any concerns they might have about receiving an unacceptable assignment. This resulted in only 21 of 1,439 (1.5%) Step-2 patients making themselves available for random assignment to all treatment options 2, p. 1235 while only 29 of 377 (7.7%) did so in Step-3. 5, p. 1521
- During each step, patients could enroll immediately into the next step if they had intolerable side-effects or had maximized their current medication(s)’ dosing without achieving a remission; and
- During any step, patients could enter continuing-care directly on their current medication(s) if they were treatment responders even if they had not achieved remission. This was done to minimize responders from dropping out in order to avoid having...
to discontinue their current medication(s) and start a new drug regimen.

Supplementary Table 2:
Description of Levels 1-4 Treatments

**Level 1:**

STAR*D investigators report that Citalopram (Celexa) was chosen as the first-line SSRI treatment because (1) absence of discontinuation symptoms; (2) demonstrated safety in elderly and medically fragile patients; (3) easy once-a-day dosing with few dose adjustments; and (4) favorable drug–drug interaction profile. Citalopram was started at 20 mg/day and then raised to 40 mg/day by day 28 and up to 60 mg/day by day 43 and onward. Dose adjustments were based on how long a patient had received a particular dose, symptom changes, and side effect burden.

**Level 2 switch treatments:**

Citalopram was discontinued without a tapering at the initiation of each level 2 switch treatment. STAR*D investigators chose pharmacologically distinct switch medications. The level 2 treatments were:

- Sertraline (Zoloft), an SSRI with the same pharmacological profile as citalopram. Sertraline was started at a daily dose of 50 mg and increased to 100 mg at day 8, to 150 mg at day 28, and to 200 mg at day 63 and onward.
- Sustained-release bupropion (Wellbutrin SR), an “out-of-class” agent whose neurochemical action mechanisms are unknown; other than that, it does not inhibit serotonin reuptake and is believed to produce antidepressant effects by blocking the reuptake of dopamine and norepinephrine. The daily dose of sustained-release bupropion was 150 mg for week 1, 200 mg from day 8 to 27, 300 mg from day 28 to 41, and 400 mg from day 42 onward.
- Extended-release venlafaxine (Effexor), a “dual-action” agent that inhibits the reuptake of both serotonin and norepinephrine. The starting daily dose of extended-release venlafaxine was 37.5 mg for week 1 and increased to 75 mg from day 8 to 14, to 150 mg from day 15 to 27, to 225 mg from day 28 to 41, to 300 mg from day 42 to 62, and to 375 mg from day 63 onward.
- Cognitive therapy was provided by a trained psychotherapist and scheduled twice weekly for the first four weeks, then once weekly for the remaining 8 weeks (16 sessions total).

**Level 2 Citalopram augmentation treatments:**

During the augmentation trial, the citalopram dose was kept constant but reduced if side effects developed. The level 2 augmentation treatments were:
- Buspirone (Buspar), a partial agonist at the postsynaptic 5-hydroxytryptamine 1A (5-HT1A) receptor that is believed to enhance the activity of SSRIs through the 5HT1A receptors. The starting dose was 15 mg per day week 1, increasing to 30 mg per day week 2, and then to 45 mg per day for weeks 3 through 5, and a final, maximum dose of 60 mg per day week 6 and onward.
- Sustained-release bupropion (Wellbutrin SR) whose neurochemical action mechanisms are unknown but is believed to produce antidepressant effects by blocking the reuptake of dopamine and norepinephrine. The initial dose was 200 mg per day during weeks 1 and 2, increasing to 300 mg per day by week 4 and to 400 mg per day week 6 and onward.
- Cognitive therapy was provided by a trained psychotherapist and scheduled twice weekly for the first four weeks, then once weekly for the remaining 8 weeks (16 sessions total).

**Level 3 switch treatments:**

At entry into the Level 3 switch trial, all level 2 medications were discontinued without tapering at the initial Level 3 treatment visit. The level 3 switch treatments were:

- Nortriptyline (Pamelor), a tricyclic antidepressant. Recommended doses were 25 mg/day for 3 days, 50 mg/day for 4 days, and then 75 mg/day by day 8, 100 mg/day by day 28, and, if necessary, 150 mg/day by day 42 and onward
- Mirtazapine (Remeron), a tetracyclic antidepressant that blocks inhibitory α2-adrenoceptors on norepinephrine and serotonin neurons to enhance both norepinephrine and serotonin neurotransmission. Recommended mirtazapine doses were 15 mg/day for the first 7 days, 30 mg/day by day 8, 45 mg/day by day 28, and, if necessary, 60 mg/day by day 42 and onward.

**Level 3 augmentation treatments of level 2 medications:**

The two medication augmentation options used in level 2, buspirone and sustained-release bupropion, were discontinued without tapering in the initial level 3 visit. The two medication augmentation treatments in level 3 were added to ongoing treatment with citalopram, sertraline, sustained-release bupropion, or extended-release venlafaxine. The level 3 augmentation treatments were:

- Lithium started at 450 mg/day, and at week 2 it was increased to the recommended dose of 900 mg/day. If participants could not tolerate the initial dose, it could be reduced to 225 mg/day for 1 week then increased to 450 mg/day. There was no monitoring of lithium levels.
- Triiodothyronine (T3), a thyroid hormone, started at 25 µg/day for 1 week and then increased to the recommended dose of 50 µg/day. There was no pretreatment assessment, nor ongoing monitoring, of thyroid functioning.
Level 4 switch treatments:

The level 4 switch treatments were:

- Tranylcypromine (Parnate), a monoamine oxidase inhibitor. A 2-week washout period of Level 3 medications was required for patients assigned to the tranylcypromine group. The recommended dosing for tranylcypromine was 10 mg/day for the first 2 weeks, followed by weekly increases of 10 mg/day until a maximum of 60 mg/day.

- Co-administered venlafaxine (Effexor) and mirtazapine (Remeron) to inhibit the reuptake of both serotonin and norepinephrine and block inhibitory 2-adrenoceptors on both norepinephrine and serotonin neurons to enhance both norepinephrine and serotonin neurotransmission. Level 3 medications were discontinued without tapering for patients assigned to this group. The dosage of extended-release venlafaxine was 37.5 mg/day for the first week, 75 mg/day for the second week, 150 mg/day for weeks 3–5, 225 mg/day for weeks 6–8, and 300 mg/day onward. Mirtazapine was started at 15 mg/day for the first 3 weeks, 30 mg/day for weeks 4 to 8, and then 45 mg/day onward.
### Supplementary Table 3:
**Number of Level 2-4 Participants Excluded from our RIAT Reanalysis, and the Reasons for their Exclusion, yet Included in STAR*D**

#### Level 2 Treatments

<table>
<thead>
<tr>
<th>Number of Level 2 Participants Excluded from our Reanalysis but Included in STAR*D</th>
<th>Bup</th>
<th>Sert</th>
<th>Ven</th>
<th>CT</th>
<th>Cit + BUP</th>
<th>Cit + Busp</th>
<th>Cit + CT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored as Remitted at <strong>ENTRY</strong> into Level 2 yet still included in STAR*D’s Level 2 analyses</td>
<td>22</td>
<td>8</td>
<td>14</td>
<td>7</td>
<td>30</td>
<td>24</td>
<td>4</td>
<td>109</td>
</tr>
<tr>
<td>Scored as only mildly depressed (<strong>HRSD &gt;7 &amp; &lt;14</strong>) at entry into Level 1, and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1, progressed to Level 2, and then included in STAR</em>D’s Level 2 data analyses</td>
<td>21</td>
<td>15</td>
<td>25</td>
<td>15</td>
<td>20</td>
<td>30</td>
<td>7</td>
<td>133</td>
</tr>
<tr>
<td>Scored as Remitted at entry into Level 1 (<strong>HRSD ≤ 7</strong>), and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1 and progressed to Level 2 and then included in STAR</em>D’s Level 2 data analyses</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>Missing baseline HRSD at entry into Level 1, and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1, and progressed to Level 2, and then included in STAR</em>D’s Level 2 data analyses</td>
<td>12</td>
<td>18</td>
<td>22</td>
<td>4</td>
<td>16</td>
<td>13</td>
<td>1</td>
<td>86</td>
</tr>
<tr>
<td>Number meeting 2 exclusion criterions</td>
<td>12</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>2</td>
<td>42</td>
</tr>
</tbody>
</table>

* Bup=Sustained-release Bupropion; Sert= Sertraline; Ven= Extended-release Venlafaxine; CT=Cognitive Therapy; Cit+BUP= Citalopram + Sustained-release Bupropion; Cit+Busp=Citalopram + Buspirone; Cit+CT= Citalopram + Cognitive Therapy

#### Level 3 Treatments

<table>
<thead>
<tr>
<th>Number of Level 3 Participants Excluded from our Data Analysis but Included in STAR*D’s Level 3 analyses</th>
<th>Nortriptyline</th>
<th>Mirtazapine</th>
<th>Lithium Augmentation</th>
<th>Triiodothyronine Augmentation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored as Remitted at <strong>ENTRY</strong> into Level 3 yet still included in STAR*D’s Level 3 analyses</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Scored as only mildly depressed (<strong>HRSD &gt;7 &amp; &lt;14</strong>) at entry into Level 1, and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1, and progressed to Level 2, and then included in STAR</em>D’s Level 2 data analyses</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>20</td>
</tr>
</tbody>
</table>
<14) at entry into Level 1, and therefore excluded from STAR*D’s data analysis, yet still treated in Level 1, progressed to Level 2, and then 3 and included in STAR*D’s Level 3 data analyses

<table>
<thead>
<tr>
<th>Scored as Remitted at entry into Level 1 (HRSD ≤ 7), and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1 and progressed to Level 2 and then 3 and included in STAR</em>D’s Level 3 data analyses</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing baseline HRSD at entry into Level 1, and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1, and progressed to Level 2, and then level 3 and included in STAR</em>D’s Level 3 data analyses</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level 4 Treatments</th>
<th>Tranylcypromine</th>
<th>Venlafaxine + Mirtazapine</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored as Remitted at ENTRY into Level 4 yet still included in STAR*D’s Level 4 analyses</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Scored as only mildly depressed (HRSD &gt;7 &amp; &lt;14) at entry into Level 1, and therefore excluded from STAR<em>D’s data analysis, yet still treated in Level 1, progressed to Level 2, 3 and then 4 and included in STAR</em>D’s Level 4 data analyses</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Scored as Remitted at entry into Level 1 (HRSD ≤ 7), and therefore excluded from STAR*D’s data analysis, yet still treated in Level 1 and progressed</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
to Level 2, then Level 3 and included in STAR*D’s Level 4 data analyses

| Missing baseline HRSD at entry into Level 1, and therefore excluded from STAR*D’s data analysis, yet still treated in Level 1, and progressed to Level 2, and then Level 3 and 4 included in STAR*D’s Level 4 data analyses | 5 | 1 | 6 |
### Supplementary Table 4:

Number and Percent of Participants Missing Entry and/or Exit HRSD Used for Last Observation Carried Forward Analyses

<table>
<thead>
<tr>
<th></th>
<th>#/(%) with Missing Entry HRSD</th>
<th>#/(%) with Missing Exit HRSD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1 (N=3,110)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 (0%)</td>
<td>926 (29.8%)</td>
</tr>
<tr>
<td><strong>Step 2 (N=1,134)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>168 (14.8%)</td>
<td>304 (26.8%)</td>
</tr>
<tr>
<td><strong>Switch strategy (N=620)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bupropion (N=190)</td>
<td>34</td>
<td>58</td>
</tr>
<tr>
<td>Sertraline (N=198)</td>
<td>27</td>
<td>56</td>
</tr>
<tr>
<td>Venlafaxine (N=192)</td>
<td>24</td>
<td>56</td>
</tr>
<tr>
<td>Cognitive Therapy (N=40)</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td><strong>Augmentation strategy (N=514)</strong></td>
<td>78 (15.2%)</td>
<td>121 (23.5%)</td>
</tr>
<tr>
<td>Bupropion (N=216)</td>
<td>35</td>
<td>58</td>
</tr>
<tr>
<td>Buspirone (N=225)</td>
<td>37</td>
<td>52</td>
</tr>
<tr>
<td>Cognitive Therapy (N=73)</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td><strong>Step 3 (N=325)</strong></td>
<td>42 (12.9%)</td>
<td>78 (24%)</td>
</tr>
<tr>
<td><strong>Level 2A (N=28)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bupropion (N=12)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Venlafaxine (N=16)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Level 3 (N=297)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Switch strategy (N=186)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nortriptyline (N=92)</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>Mirtazapine (N=94)</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td><strong>Augmentation strategy (N=111)</strong></td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>Lithium (N=58) (N=17)</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Bupropion SR (N=22)</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Citalopram (N=8)</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Sertraline (N=11)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Venlafaxine XR (N=8)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>T3 (N=53)</strong> (N=6)</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Bupropion SR (N=26)</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Citalopram (N=13)</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Sertraline (N=8)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Venlafaxine XR (N=13)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Step 4 (N=106)</td>
<td>15 (14.2%)</td>
<td>22 (20.8%)</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Level 3 (N=16)</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Tranylcypromine (N=43)</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Venlafaxine XR/mirtazapine (N=47)</td>
<td>5</td>
<td>9</td>
</tr>
</tbody>
</table>
Supplementary Table 5: Baseline Demographic and Clinical Features by Treatment Step

<table>
<thead>
<tr>
<th></th>
<th>Treatment Step</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step 1</td>
<td>Step 2</td>
<td>Step 3</td>
<td>Step 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N=3,110)</td>
<td>(N=1,134)</td>
<td>(N=325)</td>
<td>(N=106)</td>
<td></td>
</tr>
<tr>
<td><strong>Demographic Features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Mean</td>
<td>41.0</td>
<td>42.0</td>
<td>44.1</td>
<td>46.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>13.0</td>
<td>12.6</td>
<td>12.0</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>Education (years)</td>
<td>Mean</td>
<td>13.6</td>
<td>13.2</td>
<td>12.8</td>
<td>12.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3.2</td>
<td>3.3</td>
<td>3.1</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Monthly household income</td>
<td>Mean</td>
<td>2,289</td>
<td>1,744</td>
<td>1,470</td>
<td>1,003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2,732</td>
<td>1,539</td>
<td>1,383</td>
<td>887</td>
<td></td>
</tr>
<tr>
<td><strong>N</strong></td>
<td></td>
<td>1,469</td>
<td>502</td>
<td>113</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td><strong>%</strong></td>
<td></td>
<td>74.6</td>
<td>73.1</td>
<td>65.3</td>
<td>65.4</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Mean</td>
<td>2,328</td>
<td>870</td>
<td>259</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>74.9</td>
<td>76.7</td>
<td>79.7</td>
<td>81.1</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>Mean</td>
<td>333</td>
<td>115</td>
<td>29</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>10.7</td>
<td>10.1</td>
<td>8.9</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Mean</td>
<td>449</td>
<td>149</td>
<td>37</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>14.4</td>
<td>13.1</td>
<td>11.4</td>
<td>12.3</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>Mean</td>
<td>402</td>
<td>139</td>
<td>45</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>12.9</td>
<td>12.3</td>
<td>13.8</td>
<td>15.1</td>
<td></td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>Mean</td>
<td>975</td>
<td>314</td>
<td>69</td>
<td>46.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>58.7</td>
<td>54.2</td>
<td>46.9</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>Mean</td>
<td>612</td>
<td>243</td>
<td>72</td>
<td>49.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>36.9</td>
<td>42.0</td>
<td>49.0</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>Mean</td>
<td>73</td>
<td>22</td>
<td>6</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.4</td>
<td>3.8</td>
<td>4.1</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Medical insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>Mean</td>
<td>848</td>
<td>254</td>
<td>52</td>
<td>36.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>52.2</td>
<td>44.5</td>
<td>52</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Public</td>
<td>Mean</td>
<td>282</td>
<td>109</td>
<td>30</td>
<td>21.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>17.4</td>
<td>19.2</td>
<td>21.4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Mean</td>
<td>534</td>
<td>223</td>
<td>60</td>
<td>43.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>33.2</td>
<td>39.3</td>
<td>43.2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>Mean</td>
<td>475</td>
<td>171</td>
<td>40</td>
<td>27.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>28.6</td>
<td>29.5</td>
<td>27.2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>Mean</td>
<td>716</td>
<td>238</td>
<td>61</td>
<td>41.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>43.1</td>
<td>41.0</td>
<td>41.5</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Divorce/separated</td>
<td>Mean</td>
<td>429</td>
<td>155</td>
<td>42</td>
<td>28.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>25.8</td>
<td>26.7</td>
<td>28.6</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>Mean</td>
<td>41</td>
<td>16</td>
<td>4</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.5</td>
<td>2.8</td>
<td>2.7</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>Clinical Features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First episode</td>
<td>Mean</td>
<td>1,200</td>
<td>436</td>
<td>120</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>occurrence before age 18</td>
<td>SD</td>
<td>39.0</td>
<td>38.8</td>
<td>37.0</td>
<td>38.7</td>
<td></td>
</tr>
<tr>
<td>Recurrent depression</td>
<td>Mean</td>
<td>1,940</td>
<td>718</td>
<td>188</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>66.8</td>
<td>68.3</td>
<td>63.3</td>
<td>60.8</td>
<td></td>
</tr>
<tr>
<td>Family history of</td>
<td>Mean</td>
<td>1,694</td>
<td>609</td>
<td>165</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>depression</td>
<td>SD</td>
<td>55.4</td>
<td>54.9</td>
<td>51.7</td>
<td>54.7</td>
<td></td>
</tr>
<tr>
<td>Duration of current episode ≥ 2 years</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Age at first episode (years)</td>
<td>24.9</td>
<td>14.5</td>
<td>24.7</td>
<td>14.2</td>
<td>25.9</td>
<td>14.6</td>
</tr>
<tr>
<td>Illness duration (years)</td>
<td>16.1</td>
<td>13.5</td>
<td>17.2</td>
<td>13.7</td>
<td>18.2</td>
<td>14.1</td>
</tr>
<tr>
<td>Number of episodes</td>
<td>4.4</td>
<td>9.7</td>
<td>4.9</td>
<td>11.1</td>
<td>4.4</td>
<td>10.3</td>
</tr>
<tr>
<td>Duration of current episode (months)</td>
<td>25.9</td>
<td>52.0</td>
<td>28.1</td>
<td>58.8</td>
<td>32.1</td>
<td>68.5</td>
</tr>
<tr>
<td>Median duration of current episode (months)</td>
<td>8.3</td>
<td>8.7</td>
<td>9.5</td>
<td>10.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Life and Enjoyment Satisfaction Questionnaire score</td>
<td>39.1</td>
<td>14.3</td>
<td>36.5</td>
<td>13.6</td>
<td>33.7</td>
<td>13.5</td>
</tr>
<tr>
<td>SF-12 Mental c</td>
<td>25.6</td>
<td>8.1</td>
<td>25.0</td>
<td>7.7</td>
<td>24.4</td>
<td>7.7</td>
</tr>
<tr>
<td>SF-12 Physical c</td>
<td>48.6</td>
<td>12.1</td>
<td>47.0</td>
<td>12.4</td>
<td>44.5</td>
<td>12.1</td>
</tr>
<tr>
<td>Work and Social Adjustment Scale score d</td>
<td>25.0</td>
<td>8.7</td>
<td>26.3</td>
<td>8.2</td>
<td>28.3</td>
<td>7.7</td>
</tr>
<tr>
<td>HRSD_{17} score</td>
<td>21.9</td>
<td>5.2</td>
<td>22.5</td>
<td>5.2</td>
<td>23.4</td>
<td>5.2</td>
</tr>
<tr>
<td>IDS-C_{30} score e</td>
<td>39.1</td>
<td>9.6</td>
<td>40.6</td>
<td>9.7</td>
<td>42.6</td>
<td>9.4</td>
</tr>
<tr>
<td>QIDS-IVR score f</td>
<td>16.9</td>
<td>3.3</td>
<td>17.3</td>
<td>3.3</td>
<td>17.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Cumulative Illness Rating Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Categories endorsed</td>
<td>2.5</td>
<td>1.5</td>
<td>2.6</td>
<td>1.6</td>
<td>2.8</td>
<td>1.6</td>
</tr>
<tr>
<td>Total score</td>
<td>4.7</td>
<td>3.9</td>
<td>5.1</td>
<td>4.0</td>
<td>5.8</td>
<td>4.5</td>
</tr>
<tr>
<td>Severity score</td>
<td>1.8</td>
<td>0.8</td>
<td>1.8</td>
<td>0.8</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>N %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric Diagnostic Screening Questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>559</td>
<td>18.2</td>
<td>240</td>
<td>21.4</td>
<td>89</td>
<td>27.5</td>
</tr>
<tr>
<td>Alcohol abuse/dependence</td>
<td>371</td>
<td>12.0</td>
<td>136</td>
<td>12.1</td>
<td>36</td>
<td>11.1</td>
</tr>
<tr>
<td>Bulimia</td>
<td>607</td>
<td>19.7</td>
<td>232</td>
<td>20.6</td>
<td>67</td>
<td>20.7</td>
</tr>
<tr>
<td>Drug abuse/dependence</td>
<td>234</td>
<td>7.6</td>
<td>80</td>
<td>7.1</td>
<td>21</td>
<td>6.5</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>736</td>
<td>23.9</td>
<td>290</td>
<td>25.8</td>
<td>94</td>
<td>29.0</td>
</tr>
<tr>
<td>Hypochondriasis</td>
<td>336</td>
<td>10.9</td>
<td>139</td>
<td>12.4</td>
<td>45</td>
<td>13.9</td>
</tr>
<tr>
<td>Obsessive-compulsive disorder</td>
<td>723</td>
<td>23.5</td>
<td>265</td>
<td>23.6</td>
<td>97</td>
<td>29.9</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>422</td>
<td>13.7</td>
<td>183</td>
<td>16.3</td>
<td>65</td>
<td>20.1</td>
</tr>
<tr>
<td>Posttraumatic stress disorder</td>
<td>387</td>
<td>12.6</td>
<td>172</td>
<td>15.3</td>
<td>55</td>
<td>17.0</td>
</tr>
<tr>
<td>Social phobia</td>
<td>963</td>
<td>31.3</td>
<td>379</td>
<td>33.7</td>
<td>117</td>
<td>36.1</td>
</tr>
<tr>
<td>Somatoform disorder</td>
<td>284</td>
<td>9.2</td>
<td>105</td>
<td>9.3</td>
<td>35</td>
<td>10.8</td>
</tr>
<tr>
<td>Number of axis I comorbid psychiatric disorders</td>
<td>606</td>
<td>19.7</td>
<td>190</td>
<td>16.9</td>
<td>48</td>
<td>14.8</td>
</tr>
<tr>
<td>0</td>
<td>740</td>
<td>24.0</td>
<td>257</td>
<td>22.9</td>
<td>68</td>
<td>21.0</td>
</tr>
<tr>
<td>1</td>
<td>577</td>
<td>18.7</td>
<td>217</td>
<td>19.3</td>
<td>62</td>
<td>19.1</td>
</tr>
<tr>
<td>2</td>
<td>363</td>
<td>11.8</td>
<td>139</td>
<td>12.4</td>
<td>42</td>
<td>13.0</td>
</tr>
<tr>
<td>3+</td>
<td>793</td>
<td>25.8</td>
<td>321</td>
<td>28.6</td>
<td>104</td>
<td>32.1</td>
</tr>
</tbody>
</table>

* Sums do not always equal N due to missing values. Percentages are based on available data.

* Integrated voice response (IVR) administered version of the Quality-of-Life Enjoyment and Satisfaction Questionnaire assessing participants’ global rate of satisfaction. Higher scores (range=0–100) represent greater life enjoyment and satisfaction.

* IVR-administered version of the SF-12 assessing perceived mental and physical health status. Two subscales (physical health factor and mental health) range from 0 to 100—higher scores indicate better functioning with a population norm for each score of 50.

* IVR-administered version of the Work and Social Adjustment Scale. Scores between 10 and 20 are associated with significant functional impairment while scores above 20 suggest moderate to severe functional impairment.

* Inventory of Depressive Symptomatology administered telephonically.

* IVR-administered version of the QIDS.