Patient-reported outcomes before and after treatment of major depressive disorder

Waguih William IsHak, MD, FAPA; James Mirocha, MS; Sarah Pi, BS; Gabriel Tobia, MD; Bret Becker, MD, MS; Eric D. Peselow, MD; Robert M. Cohen, PhD, MD

Introduction

Major depressive disorder (MDD) is one of the leading causes of disability and premature mortality, and is predicted to become the second most burdensome condition worldwide by the year 2020.[1-3] Patient-reported outcomes (PROs) are increasingly utilized

*Patient reported outcomes (PROs) of quality of life (QoL), functioning, and depressive symptom severity are important in assessing the burden of illness of major depressive disorder (MDD) and to evaluate the impact of treatment. We sought to provide a detailed analysis of PROs before and after treatment of MDD from the large Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. This analysis examines PROs before and after treatment in the second level of STAR*D. The complete data on QoL, functioning, and depressive symptom severity, were analyzed for each STAR*D level 2 treatment. PROs of QoL, functioning, and depressive symptom severity showed substantial impairments after failing a selective serotonin reuptake inhibitor trial using citalopram (level 1). The seven therapeutic options in level 2 had positive statistically (P values) and clinically (Cohen's standardized differences [Cohen's d]) significant impact on QoL, functioning, depressive symptom severity, and reduction in calculated burden of illness. There were no statistically significant differences between the interventions. However, a substantial proportion of patients still suffered from patient-reported QoL and functioning impairment after treatment, an effect that was more pronounced in nonremitters. PROs are crucial in understanding the impact of MDD and in examining the effects of treatment interventions, both in research and clinical settings.*

Dialogues Clin Neurosci. 2014;16:171-183.

Keywords: *major depressive disorder; quality of life; functioning; individual burden of illness; STAR*D*

Author affiliations: Vice Chairman for Education & Research and Consultant Psychiatrist, Cedars-Sinai Department of Psychiatry and Behavioral Neurosciences, Los Angeles, California, USA (Waguih William IsHak); Cedars-Sinai Department of Psychiatry and Behavioral Neurosciences, Los Angeles, California, USA (Waguih William IsHak, James Mirocha, Sarah Pi, Gabriel Tobia, and Bret Becker); David Geffen School of Medicine at University of California Los Angeles (UCLA), Los Angeles, California, USA (Waguih William IsHak); Richmond University Medical Center and Freedom From Fear, Staten Island, New York, USA (Eric D. Peselow); Department of Psychiatry, Emory University School of Medicine (Robert M. Cohen)

Address for correspondence: Waguih William IsHak, MD, FAPA, Vice Chairman for Education & Research and Consultant Psychiatrist, Department of Psychiatry and Behavioral Neurosciences, Cedars-Sinai Medical Center, 8730 Alden Drive, Thalians E-132, Los Angeles, CA 90048, USA (e-mail: waguih.ishak@cshs.org)

to assess patients with MDD and consist of self-rating scales of symptom severity, functioning, and quality of life (QoL). PROs are being adopted, developed, and promoted by the World Health Organization (WHO) International Classification of Functioning, Disability, and Health (ICF),[4] US National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS),[5] US Patient-Centered Outcomes Research Institute (PCORI),[6] US Federal Food and Drug Administration (FDA)–supported initiative for PROs (Critical Path Initiative [CPI]),[7] and UK National Health Service (NHS) Patient-Reported Outcome Measures (PROMs).[8]

Depressive symptoms may affect patients' self-reports of symptom severity, functioning, and QoL,[9] with their influence being shown as a mediator variable.[10] Similarly, poor QoL, functional impairment and more severe symptoms could also result in worsening of depression.[11] This bidirectional relationship continues to interfere with the precision of outcome measurements. Although widely used, clinician-rated measures of symptom severity are not immune from patient bias since they are primarily based on patient reports in addition to clinician observation, and they were even reported to yield a significant discrepancy when administered to patients vs their informants.[12] Except in dysthymic and nonendogenous depressed groups, empirically designed self-report scales tend to have a moderately high correlation with clinician-rated ones.[13] Moreover, PROs continue to provide valuable information that could not be obtained using clinician-rated measures despite the risk of minimization or magnification of the actual burden of illness. In fact, QoL by conceptual and operational definitions has to be measured by subjective reporting. Based on the WHO definitions, QoL reflects the patient's satisfaction with health and life activities, ie, work, love, and play activities by self-report,[14] whereas functioning refers to an individual's actual involvement and participation in the aforementioned activities as rated by self or observers.[15] Unless clinicians are using collateral information, functioning is primarily measured by self-rating. Individuals with MDD frequently suffer from QoL and functioning impairments,[16] and several investigators have demonstrated that treatment of severe mental disorders, including MDD, should not only focus on reduction of symptoms, but also seek to enhance levels of functioning, and more significantly improve the patient's subjective wellbeing and QoL.[17,18] Since patients remain at the center of suffering from depression, their perceptions of the dimensions of their burden of illness using PROs should remain as fundamental tools for assessment of the effectiveness of treatment interventions, especially in the era of patient-centered health care. Descriptions of the most commonly used PROs in MDD appear in *Table I*.[19-33]

In order to present a real-world application of PROs, we are seeking to provide a detailed analysis of PROs before and after treatment of MDD using different interventions following the failure of first-line treatment with a selective serotonin reuptake inhibitor (SSRI). In earlier studies, we examined the effect of level 1 treatment on functioning, QoL, and the individual burden of illness for depression (IBI-D), a vector derived from a principal component analysis that captures the vast majority of the variance in PROs of depressive symptom severity, functioning, and QoL in depressed patients.[34] The present analysis aims at examining PROs in patients who failed first-line treatment with the SSRI citalopram, before and after enrolment in level 2 using seven interventions as second-line interventions in MDD. Although many publications[35-38] have reported findings from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, patient-reported functional outcomes were seldom described. Moreover, previous publications have traditionally focused on results based on *P* values, whereas the current analysis adds the examination of effect sizes in order to assess more clinically meaningful effects of the interventions as assessed using patient-reported measures.

Methods

Study population

STAR*D is the largest study ever conducted on MDD treatment, and featured multiple levels of pharmacotherapy and psychotherapy trials. Patients failing one trial of SSRI monotherapy (level 1), were either switched from citalopram to sertraline, venlafaxine, bupropion, or cognitive therapy, or kept on citalopram and augmented with bupropion, buspirone, or cognitive therapy (level 2). The STAR*D study was funded by the National Institute of Mental Health (NIMH) and is the largest study aimed at analyzing subsequent treatment steps for patients with treatment-resistant MDD. The design and rationale of STAR*D are detailed elsewhere.[39,40] The study enrolled 4041 18- to 75-year-old outpatients with nonpsychotic

Name	Time to complete	Number of items	Item scale	Score range	Higher score is	Summary
QoL patient-reported outcomes						
Q-LES-Q Quality of Life, Enjoyment, and Satisfaction Questionnaire—Short Form	5 min	16	1-5	0-100	Better	The Q-LES-Q assesses QoL covering the following domains: health, mood, work, household activities, social and family relationships, leisure, ability to function daily, sex, economic and living situation, mobility, vision, wellbeing, medications, and overall satisfaction. The total score is calculated as the sum of scores from items 1 through 14 and is converted to a percentage using the following calculation: ([raw score-14]/56) × 100. Community norms scores had a mean of 78.3%.[19]
SF-36 Medical outcomes study—Short Form 36	15 min	36	0-5	0-100	Better	The SF-36 and its brief form the SF-12 measure QoL on eight health concepts: 1. Limitations in physical activities because of health problems, 2. Limitations in social activities because of physical or emotional problems, 3. Limitations in usual role activities because of physical health problems, 4. Bodily pain. 5. General mental health (psychological distress and well-being).
SF-12 Medical outcomes study—Short Form 12	5 min	12				
WHOQOL WHO Quality of Life	25 min	100	1-5	0-100	Better	The WHOQOL and its brief form the WHOQOL-BRF are focused around the definition of QoL advocated by WHO; this includes the culture and context that influence an individual's perception of health. They measure four domains: physical health, psychological health, social relationships, and environment.[21]
WHOQOL-BREF 26-item version	10 min	26				
EQ-5D EuroQoL	3 min	5	1-3	-1.0 to 1.0	Better	The EQ-5D measures QoL using five single-item measures of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Scores range from 1.0 (perfect health) to -1.0 (death). It has an additional visual analog scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).[22]
PROMIS-GHS Patient-Reported Outcomes Measurement Information System Global Health Scale	5 min	10	0-3	0-100	Better	The PROMIS-GHS measures health and QoL by assessing five primary domains: physical function, fatigue, pain, emotional distress, and social health. Scoring results in a "physical health" component and a "mental health" component each with a mean of 50 (SD, 10), where higher or lower scores indicate better or worse health than the population.[23]
Functioning patient-reported outcomes						
WSAS Work and Social Adjustment Scale	3 min	5	0-8	0-40	Worse	The WSAS measures functioning in the work, home management, private leisure, social leisure, and relationship domains. The sum of the scores produces a total score where a score >20 indicates major functional impairment, 10-20 indicates significant functional impairment, and scores <10 are within normal range.[24]

Table I. Patient-reported outcomes of quality of life (QoL), functioning, and depressive symptom severity. DSM-IV, Diagnostic and Statistical Manual of Mental Disorders—IV; MDD, major depressive disorder; SD, standard deviation; WHO, world Health Organization

Name	Time to complete	Number of items	Item scale	Score range	Higher score is	Summary
WHODAS 2.0 WHO Disability Assessment Schedule 2.0	15 min	36	0-3	0-100	Worse	The WHODAS 2.0 and its brief 12-item version measure functioning in: cognition (understanding and communicating); mobility (moving and getting around); self-care (hygiene, dressing, eating, and staying alone); getting along (interacting with other people); life activities (domestic responsibilities, leisure, work, and school); and participation (joining in community activities). Scoring utilizes one of two methods: simple scoring involves simple sum of the score, and complex scoring uses a script converting the score using item-response theory to a range from 0 (no disability) to 100 (total disability).[25]
WHODAS 2.0 12-item version	5 min	12				
SDS Sheehan Disability Scale	3 min	3	0-10	0-30	Worse	The SDS assesses functioning in the domains of work, social life, and family life/home responsibilities. The sum of the scores lead to a total score ranging from 0 (unimpaired) to 30 (highly impaired). Scores ≥5 on any of the domains or total score ≥8 are indicative of functional impairment.[26]
EWPS Endicott Work Productivity Scale	10 min	25	0-4	0-100	Worse	The EWPS covers twenty-five aspects of work/job functions such as being on time, accomplishing tasks, and performance. The item scores are summed up to a total score that ranges from 0 (no impairment) to 100 (major impairment in work productivity).[27]
Depressive symptom severity patient-reported outcomes						
QIDS-SR Quick Inventory of Depressive Symptomatology—Self Report	10 min	16	0-3	0-27	Worse	The QIDS-SR measures the severity of 16 depressive symptoms. The total score is a sum of the highest score on any one of four sleep items (1-4) + item (5) + the highest score on any one appetite/weight item (6-9) + items (10-14) + the highest score on either of the two psychomotor items (15 and 16). Severity of MDD depressive symptoms is categorized based on the QIDS-SR scores: 0-5 (remission), 6-10 (mild), 11-15 (moderate), 16-20 (severe), or >20 (very severe).[28]
BDI-II Beck Depression Inventory II	10 min	21	0-	0-6	Worse	The BDI-II measures the severity of 21 depressive symptoms. The total score is the sum of all items. Depression severity is categorized with scores of 0-13 (minimal depression), 14-19 (mild depression), 20-28 (moderate depression), 29-63 (severe depression).[29]
CUDOS Clinically Useful Depression Outcome Scale	5 min	18	0-4	0-64	Worse	The CUDOS rates 16 *DSM-IV* depression symptoms from "not at all true" (0 days) to "almost always true" (every day), item 17 rating interference with functioning, and item 18 rating quality of life. The total score is the sum of the first 16 items, ranging from 0-10 (nondepressed), 11-20 (minimal depression), 21-30 (mild depression), 31-45 (moderate depression), or 46 and above (severe depression).[30]
CES-D Center for Epidemiologic Studies Depression Scale	10 min	20	0-3	0-60	Worse	The CES-D measures the severity of 20 depressive symptoms from "rarely" to "most of the time". The score is the sum of the 20 questions. A score of 16 points or more is considered as "depressed".[31]

Table I. Continued

MDD from 18 primary care and 23 psychiatric care practice settings across the United States from 2001 to 2007. There are up to four sequential treatment levels, but this study focuses only on level 2, which includes seven treatment options. Patients from level 1 who were unable to tolerate citalopram, as well as those who failed to show adequate improvement, were offered three augmentation options (adding sustained-release bupropion, buspirone, or cognitive therapy in addition to the level 1 citalopram) and four switch options (replacing citalopram with sertraline, sustained-release bupropion, extended-release venlafaxine, or cognitive therapy). The authors obtained a NIMH data use certificate to access and use the STAR*D Pub Ver1 dataset for this analysis. The complete data on QoL, functioning, and depressive symptom severity was analyzed for each level 2 treatment.

Outcome measures

Quality of Life, Enjoyment, and Satisfaction Questionnaire–Short Form (Q-LES-Q-SF)

The Q-LES-Q-SF,[19] used to assess QoL, is a self-report questionnaire that measures satisfaction and enjoyment in a series of discrete domains and life activities. This study uses the short version, which has 16 items: physical health, mood, work, household activities, social relationships, family relationships, leisure time activities, ability to function in daily life, sexual drive/interest/performance, economic status, living/housing situation, ability to get around physically, vision, and overall sense of wellbeing. Each item is scored on a 5-point Likert scale, where 1=very poor, 2=poor, 3=fair, 4=good, and 5=very good. Adding the results of the 14 first items, then dividing by the maximum score, and multiplying this figure by 100 gives a total score ranging from 0 to 100, with 0 being the lowest QoL score and 100 the highest. Community norm samples have a mean Q-LES-Q-SF score of 78.3 (standard deviation [SD], 11.3) and scores within 10% of this value, ie, Q-LES-Q-SF ≥70.47, are considered "normal", whereas Q-LES-Q-SF scores greater than 2 SD below the community norm scores, ie, Q-LES-Q-SF ≤55.7, are considered "severely impaired."[19] The Q-LES-Q-SF has a Cronbach α of 0.90 and test-retest reliability of 0.74, demonstrating strong psychometric properties.[19]

Work and Social Adjust Scale (WSAS)

The WSAS[41] was used to assess functioning. The WSAS is a five-item self-report questionnaire that measures impairment of functioning in various settings: work; home management, eg, cleaning, tidying, shopping, cooking, looking after home or children, and paying bills; social leisure time, ie, activities done with other people, such as parties, clubs, outings, dating, and home entertainment; private leisure time, ie, activities done alone, such as reading, gardening, collecting, sewing, and walking alone; and ability to form and maintain close relationships with others. Each item is scored on a visual analogue scale ranging from 0 (no impairment at all), to 8 (very severe impairment). The sum of the results of the five items gives a total score ranging from 0 (best functioning) to 40 (worst functioning). Severe impairment is indicated by scores above 20.[41] The WSAS has a Cronbach α varying from 0.70 to 0.94 and test-retest reliability of 0.73, also demonstrating strong psychometric properties.[41]

Name	Time to complete	Number of items	Item scale	Score range	Higher score is	Summary
PHQ-9 Patient Health Questionnaire	5 min	9	0-3	0-27	Worse	The PHQ-9 measures the nine depressive symptoms from the *DSM-IV*. The total score is the sum of the nine items with scores of 1-4 (minimal depression), 5-9 (mild depression), 10-14 (moderate depression), 15-19 (moderately severe depression), or 20-27 (severe depression).[32]
PROMIS Depression Patient-Reported Outcomes Measurement Information System Depression Scale	5 min	8	1-5	0-100	Better	The PROMIS Depression scale measures negative mood, view of self, social cognition, decreased positive affect, and engagement. The raw score is then converted to a T score that has a population mean of 50 (SD, 10).[33]

Table I. Continued

Quick Inventory of Depressive Symptomatology—Self Report (QIDS-SR)

The QIDS-SR[28] was used to assess severity of depressive symptoms. The QIDS-SR is a 16-item questionnaire corresponding to the nine *DSM-IV* criteria of major depression: one item for each of the following symptoms: depressed mood, decreased interest, decreased energy, worthlessness/guilt, concentration/decision making, and suicidal ideation; four items to assess sleep: early, middle, and late insomnia, and hypersomnia; two items to assess psychomotor disturbance: agitation and retardation; and four items to assess appetite/weight: appetite increase or decrease, and weight increase or decrease. Each item is rated 0 to 3. The QIDS-SR score is calculated by summing the scores of the items. In domains utilizing more than one item (eg, the four items for sleep disturbance), only the highest score is utilized in the total score.[28] The QIDS-SR scores range from 0 (not depressed) to 27 (most severely depressed). A score of five or less indicates remission, which is the goal of treatment.[28] The QIDS-SR has high internal consistency (Cronbach α, 0.86) and is highly associated with the three versions of the clinician-rated Hamilton Rating Scale for Depression, Montgomery-Åsberg Depression Rating Scale, and the Beck Depression Inventory.[28]

Individual Burden of Illness for Depression (IBI-D)

The burden of illness was measured using the IBI-D, a newly introduced measure that incorporates QoL, functioning, and depressive symptom severity.[34] The IBI-D is the first and only statistically significant principal component obtained from a principal component analysis of the above three well-validated PROs of depressive symptom severity (QIDS-SR), functioning (WSAS), and QoL (Q-LES-Q-SF). IBI-D is a z-score that references patients in level 1 of STAR*D, where values around 0 represent a burden similar to the average depressed patient, a burden greater than +2 indicates that the patient has an unusually high burden (higher than the top 2% of depressed patients), and values lower than -2 indicate that the patient has a lower burden (lower than 98% of depressed patients). The IBI-D was shown to adequately capture the multidimensional impact of antidepressant treatment,[42] and to adequately predict relapse in MDD.[43]

Statistical methods

Summary values are expressed as means and SDs for continuous variables, and frequencies (%) for categorical variables. We calculated effect sizes using Cohen's standardized differences (Cohen's *d*), in order to assess clinical significance in addition to statistical significance. Cohen's *d* values of 0.2, 0.5, and 0.8 describe small, medium, and large effects, respectively. Within the treatment groups, changes from entry to exit (Δs) on continuous variables were assessed for significance using paired *t* tests. Between-group differences on continuous variables were assessed for significance using independent sample *t* tests. We calculated and compared the proportions of patients with normal QoL (using Q-LES-Q-SF) and functioning (using WSAS) and with severe impairments on both measures. Within the treatment groups, preintervention vs postintervention *P* values were calculated using McNemar's test for related proportions. Five tests were performed for each outcome measure: two within-group tests and three between-group tests (entry, change, and exit). Thus, we used a Bonferroni-adjusted 0.01 significance level for each test. Analyses were performed using SAS software version 9.2 (SAS Institute Inc, Cary, NC, USA).

Results

STAR*D used an equipoise stratified randomized design.[44] In level 2, patients with the help of their clinicians considered either switching or augmenting citalopram using a new medication or cognitive therapy. Patients could decline any strategy, however there had to be at least two possible strategies, to one of which the patient was randomized.[45]

Study population demographics

Complete QoL, functioning, and depressive symptom severity data from STAR*D level 2 trial subjects were analyzed (n=749). The mean age was 44.4 years (SD, 12.4) with a range from 18.8 to 75 years. Caucasians represented 81.6% (n=611), while Hispanics accounted for 12.3% (n=92) in this sample. Females made up 60.1% (n=450) and 25.1% (n=188) were college graduates. Slightly more than half, 51.9% (n=389), were employed at the start of the study and 42.1% (n=330) were living with a spouse or partner.

QoL, functioning, and depressive symptom severity scores

Level 2 STAR*D entry and exit scores for QoL (Q-LES-Q-SF), functioning (WSAS), and depressive symptom severity (QIDS-SR) were analyzed for the each of the seven treatment options separately to see each option's effect in MDD.

Pretreatment scores

The Q-LES-Q-SF scores ranged from 38.5 to 45.2 with a mean score of 42.1 (SD, 15.5). The WSAS scores ranged from 22.4 to 24.1 with a mean score of 23.6 (SD, 9.0). The QIDS-SR scores ranged from 13.1 to 15.7 with a mean score of 14.3 (SD, 4.7). No statistically significant differences were found between the seven treatment groups.

Post-treatment scores: impact of treatment on QoL, functioning, and depressive symptoms

All seven treatment options led to significant improvements in patients, both statistically (P<0.01) and clinically (Cohen's d, >0.4). The treatment with the highest effect size for QoL, as measured by Q-LES-Q-SF, was switching to cognitive therapy (Cohen's d, 0.73) followed by switching to sertraline (Cohen's d, 0.67). The lowest effect sizes were for switching to bupropion and augmenting citalopram with cognitive therapy (Cohen's d, 0.42). For functioning, as measured by WSAS, the highest effect sizes were with switch to cognitive therapy (Cohen's d, 0.78), followed by sertraline, venlafaxine, and citalopram plus bupropion (Cohen's d, 0.62), and the lowest was bupropion alone (Cohen's d, 0.47). The highest effect sizes for depressive symptoms severity, as measured by QIDS-SR, were for venlafaxine (Cohen's d, 0.88) and cognitive therapy (Cohen's d, 0.83), with the lowest being citalopram plus buspirone (Cohen's d, 0.48). It is especially important to note the high effect sizes for venlafaxine and cognitive therapy (Cohen's d, >0.8), indicating large, clinically significant improvements. In general, the effect sizes were higher for patient-reported depressive symptom severity, than for QoL or functioning.

Proportions of patients scoring within normal QoL and functioning

Level 2 STAR*D entry and exit proportions of patients with normal QoL (Q-LES-Q-SF ≥78.3) and functioning are presented in *Table II*.[19,46,47]

Pretreatment proportions

Before treatment, depressed patients scoring for a normal QoL ranged from 0% to 5%, with a mean proportion of 2.9%. Patients within normal functioning ranged from 1.7% to 11.5%, with a mean proportion of 6.4%.

Post-treatment proportions: impact of treatment

The posttreatment data revealed a statistically significant increase in the proportion of patients with normal QoL and functioning. All P values were less than 0.0005 for QoL and functioning, except for citalopram plus cognitive therapy (QoL, P=0.63; functioning, P=0.012). However, the proportion of patients with normal QoL and functioning upon exit ranged from Q-LES-Q-SF scores of 6.4% to 31.3%, and WSAS scores of 21.3% to 43.8%, with the mean percentage of patients with normal QoL of 19.5%, and with normal functioning of 31.9%. In other words, following treatment, only 1 out of 5 patients achieved a normal QoL and less than one third of patients achieved normal functioning.

Proportions of patients with severe impairments in QoL and functioning

Level 2 STAR*D entry and exit proportions of patients with severe impairments in QoL (2 SD below community norms, ie, Q-LES-Q-SF <55.7) and functioning (WSAS >20) are displayed in *Table II*.[19,46,47]

Pretreatment proportions

Before treatment, severe impairments in QoL and functioning were detected in a large proportion of patients, with an overall higher percentage of patients with severely impaired QoL (range, 77.2% to 93.6%; mean, 83.3%) than functioning (range, 57.7% to 66.1%; mean, 62.5%), ie, nearly 4 out of 5 patients suffer from severe QoL impairment and 2 out of 3 patients suffer from severe functional impairment.

Post-treatment proportions: impact of treatment

The post-treatment data revealed a statistically significant reduction in the proportion of patients suffering from severely impaired QoL ($P<0.005$). Several of the functional impairment results were not statistically significant at the Bonferroni-adjusted 0.01 significance level. The proportion of patients severely impaired in QoL and functioning upon exit ranged from 68.6% to 43.8% (mean, 59.5%) and from 49.6% to 37.5% (mean, 44.3%), respectively, with a mean of 59.5% (QoL) and of 44.3% (functioning) for the whole sample. In other words, nearly 2 out of 3 patients still struggled with severe impairments in QoL, and less than half still experienced severe impairments in functioning. Nonremitters had substantially smaller proportions of patients within normal QoL and functioning, and larger proportions with severe impairments in QoL and functioning at exit.

Individual burden of illness for depression (IBI-D) scores

Level 2 STAR*D entry and exit scores for burden of illness for depression (IBI-D) are displayed in *Table III*.

Pretreatment scores

Generally, the baseline IBI-D scores of remitters were lower than that of nonremitters, suggesting that remitters started out with less burden of illness overall. For remitters ($P<0.0001$), the treatment group with the most burden at baseline was cognitive therapy (z score, -0.13),

Intervention	n	Within normal QoL Pre (%)	Within normal QoL Post (%)	Mcnemar Test *P* value	Severely impaired QoL Pre (%)	Severely impaired QoL Post (%)	Mcnemar test *P* value
Bupropion	121	5.0	20.7	0.0002	84.3	68.6	0.002
Sertraline	131	0	17.6	<0.0001	89.1	62.6	<0.0001
Venlafaxine	121	1.7	14.0	0.0003	89.3	61.2	<0.0001
Cognitve therapy	32	3.1	31.3	0.012	84.4	43.8	0.001
Citalopram+bupropion	148	4.1	23.6	<0.0001	75.0	53.4	<0.0001
Citalopram +buspirone	149	4.0	22.1	<0.0001	77.2	55.0	<0.0001
Citalopram +cognitve therapy	47	2.1	6.4	0.63	93.6	68.1	0.004
All	749	2.9	19.5	<0.0001	83.3	59.5	<0.0001
Intervention	**n**	**Within normal functioning Pre (%)**	**Within normal functioning Post (%)**	**Mcnemar Test *P* value**	**Severely impaired functioning Pre (%)**	**Severely impaired functioning Post (%)**	**Mcnemar test *P* value**
Bupropion	121	6.6	25.6	<0.0001	64.5	47.9	0.0008
Sertraline	131	5.3	31.3	<0.0001	65.6	49.6	0.002
Venlafaxine	121	1.7	28.1	<0.0001	66.1	45.5	0.0001
Cognitve therapy	32	3.1	43.8	0.0002	65.6	37.5	0.012
Citalopram +bupropion	148	11.5	39.2	<0.0001	58.8	41.9	0.0003
Citalopram +buspirone	149	8.1	34.2	<0.0001	57.7	39.6	<0.0001
Citalopram +cognitive therapy	47	2.1	21.3	0.012	66.0	44.7	0.031
All	749	6.4	31.9	<0.0001	62.6	44.3	<0.0001

Table II. Proportions of patients with normal quality of life (QoL) and functioning before and after each intervention. Normal QoL is defined as Q-LES-Q-SF scores within 10% of community norms, and severe impairment is defined as Q-LES-Q-SF scores greater than 2 standard deviations (SD) below the community norms. Since community norm samples have an average Q-LES-Q-SF of 78.3 (SD, 11.3), a Q-LES-Q-SF ≥70.47 is considered within normal and a Q-LES-Q-SF ≤55.7 is considered severely impaired.[19,46,47] Normal functioning is defined as WSAS scores of less than 10 and severe impairment is defined as WSAS scores of more than 20.[25] n, number

while the treatment group with the least burden was citalopram plus bupropion (*z* score, -0.77) followed by citalopram plus buspirone (*z* score, -0.76). For nonremitters (all *P* values, <0.0001), the treatment group with the most burden was sertraline (*z* score, 0.25) followed by venlafaxine (*z* score, 0.24), while the treatment group with the least was citalopram plus cognitive therapy (*z* score, -0.10) followed by cognitive therapy (*z* score, -0.11) and citalopram plus bupropion (*z* score, -0.11).

Post-treatment scores: impact of treatment

Posttreatment, the data revealed an overall statistically significant reduction in the burden of illness for depression (all *P* values, <0.0001). For both remitters and nonremitters (all *P* values, <0.0001), the treatment group with the least burden upon exit was cognitive therapy (change for IBI-D scores in remitters from -0.13 to -2.68; nonremitters from -0.11 to -0.50). Additionally, the treatments that led to the greatest decrease in burden (all *P* values, <0.0001) were cognitive therapy (change for IBI-D scores in remitters, -2.55; nonremitters, -0.39), sertraline (change for IBI-D scores in remitters, -1.96; nonremitters, -0.42), and venlafaxine (change for IBI-D scores in remitters, -1.90; nonremitters, -0.49).

Differences between the seven interventions on PROs

There were no statistically significant differences between the interventions. Interestingly, switching to cognitive therapy stood out numerically with the greatest effect size for QoL (Cohen's *d*, 0.73), the largest proportion of patients with normal QoL and functioning upon exit (31.3%), the lowest proportion of patients severely impaired (43.8%), as well as the lowest IBI-D score upon exit (*z* score, -0.50).

Discussion

The main findings of the present study are: (i) PROs show that the seven level 2 treatment options produced significant functional improvements that were significant both statistically (*P*<0.01) and clinically (Cohen's *d*>0.4); and (ii) patient-reported functional outcomes revealed that a substantial proportion of patients who had failed a first-line trial with citalopram, still experienced grave impairments in QoL and functioning after treatment with second-line augmentation or switching interventions, an effect that is more pronounced in nonremitters.

The seven different treatments of MDD in patients who have failed initial citalopram monotherapy show a significant positive impact on QoL, functioning, and depressive symptom severity using patient-reported measures. The use of effect size enabled us to assess the magnitude of this impact after ascertaining that it did not happen by chance, ie, after establishing statistical significance. On depressive symptom severity, other STAR*D analyses[36,38] concluded that bupropion, sertraline, venlafaxine, and cognitive therapy, as well as citalopram plus bupropion, citalopram plus buspirone, and citalopram plus cognitive therapy, lead to similar outcomes. However, this study is unique in including patient-reported functional outcomes. In general, the effect sizes were the highest for patient-reported depressive symptom severity, followed by functioning, then QoL. Switching to cognitive therapy alone, after failing first-line SSRI treatment, achieved numerical superiority, which is a finding that might be worthy of future exploration. Although the usefulness of cognitive therapy in refractive MDD has long been established[48] with respect to reducing symptoms, the present study showed that the cognitive therapy group displayed the largest proportion of patients with normal QoL and functioning upon exit, the smallest proportion of patients severely impaired, the greatest effect size for both QoL and functioning, and also displayed the lowest IBI-D score upon exit. Previous STAR*D analyses utilized remission (not patient-reported QoL or functioning) as primary outcomes to show that cognitive therapy alone, or in addition to citalopram, was as effective as the other level 2 pharmacologic options with "comparable outcomes."[36] The above results for cognitive therapy need to be interpreted with extreme caution; the results should only be considered as hypothesis-generating, since STAR*D was not designed to compare it with other treatments and the sample size was small.

Although the impact of the seven interventions is statistically and clinically significant, a substantial proportion of patients failed to achieve normal scores on QoL and functional patient-reported measures. Nonremitters showed remarkably large proportions with severe impairments in QoL and functioning. We also observed considerably small proportions of nonremitters who experienced normal QoL and functioning. Moreover, our analysis shows that, even after achieving remission, only half of the remitters scored a normal QoL. This finding adds more credence to the notion that remission (mini-

mal or no symptoms) does not reflect full recovery of functional outcomes, and perhaps more ways to improve QoL and functioning will need to be researched and applied in clinical settings.[18]

Strengths and limitations

Our study suffers from a number of limitations, related both to our own analysis and to the STAR*D study design. It is important to note the original STAR*D study was not designed to accommodate comparison of level 2 interventions. It could be speculated that the lack of equipoise in patient randomization affected statistical power and, down the line, clinical significance. It would be useful to design another experiment specifically for this purpose, with a more balanced distribution of patients. Additionally, it would be informative to compare cognitive therapy to other forms of psychotherapy that have already proven their usefulness in treating MDD, such as interpersonal psychotherapy.[49]

STAR*D lacked a control or a placebo group, which could have provided useful comparative data and helped control for the placebo effect and the passage of time as factors of remission. However, this may be of less significance for level 2 than for level 1 data. No blinding was required for the physicians and the patients involved. There is also a dearth of information regarding dropouts. Warden et al[50] have demonstrated that African-Americans, young patients, individuals with lower education, and patients with lower income were more likely to drop out of the STAR*D study than other groups.[50] The vast majority (90%) of cases of attrition in level 2 of the STAR*D study were shown to be motivated by nonmedical reasons.[51] This finding might explain why no difference was found in the percentage of patients that exited the study in the drug group compared with the cognitive therapy group, which lacked a drug side effect.[36] Medical predictors of attrition included medication side effects and the presence of *Diagnostic and Statistical Manual of Mental Disorders (DSM)* axis I comorbidities.[50] Attrition makes it more difficult to extrapolate conclusions from the sample studied to the general population; it is therefore important for future studies to account for missing data from attrition to avoid selection bias. Another potential weakness in this present analysis is the lack of control for coexisting symptoms, such as anxiety, insomnia, and loss of energy. Gaynes et al[52] have suggested that, while the latter do affect remission rates, symptoms such as loss of energy may guide medication selection according to side-effect profile.[52]

One of the most important strengths of this present study is its examination of effect sizes. While STAR*D studies have traditionally analyzed results using *P* values, the current analysis uses Cohen's *d* in order to compare the relative strengths of the seven interventions. This estimation of magnitude serves to complement the statistical inference supplied by *P* values.

Another important strength was the reliance on PROs. The instruments have already demonstrated strong psychometric properties, and provided unique perspective on functional outcomes that are difficult to obtain by clinician rating. PROs were identified to play an important role not only in examining the impact of treatment interventions, but also in predicting relapse in MDD.[53]

	All					Remitters		
Intervention	**n**	**IBI-D Pre**	**IBI-D Post**	**IBI-D Change**	***P***	**n**	**IBI-D Pre**	
Bupropion	121	-0.02 (1.07)	-0.67 (1.44)	-0.66 (1.21)	<0.0001	26	-0.58 (1.03)	
Sertraline	131	0.04 (0.95)	-0.79 (1.43)	-0.83 (1.15)	<0.0001	35	-0.54 (0.93)	
Venlafaxine	121	0.08 (0.92)	-0.81 (1.35)	-0.89 (1.08)	<0.0001	34	-0.32 (0.79)	
Cognitive therapy	32	-0.12 (0.85)	-1.32 (1.41)	-1.20 (1.46)	<0.0001	12	-0.13 (0.94)	
Citalopram + Bupropion	148	-0.34 (0.94)	-1.14 (1.35)	-0.80 (1.09)	<0.0001	53	-0.77 (0.77)	
Citalopram + Buspirone	149	-0.30 (0.98)	-0.90 (1.42)	-0.59 (1.07)	<0.0001	38	-0.76 (0.72)	
Citalopram + Cognitive therapy	47	-0.14 (0.86)	-0.86 (1.09)	-0.72 (1.14)	<0.0001	10	-0.32 (0.73)	
All	749							

Table III. Comparisons for 749 patients with both pre- and post-treatment values using the Individual Burden of Illness for Depression (IBI-D). Values are means (standard deviation [SD]). Paired t test (within intervention change: exit vs base). *P*=0.082 (nonsignificant [NS]) for difference in base to exit change between interventions, confirmed by Welch's analysis of variance (ANOVA; *P*=0.143; NS). n, number

Conclusion

PROs showed that all seven treatment options led to statistically and clinically significant improvements, however a substantial proportion of patients still suffered from QoL and functional impairments. Despite remission, only half of the remitters scored for a normal QoL. PROs play an important role in examining treatment interventions in MDD, both in research and clinical settings. ❑

REFERENCES

1. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet*. 1997;349:1498-1504.
2. Kessler RC, Chiu WT, Demler O, et al. Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Arch Gen Psychiatry*. 2005;62:617-627.
3. Zung WW, Broadhead WE, Roth ME. Prevalence of depressive symptoms in primary care. *J Fam Pract*. 1993;37:337-344.
4. World Health Organization. International Classification of Functioning, Disability and Health (ICF). Available at: http://www.who.int/classifications/icf/en/. Accessed February 11, 2014.
5. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care*. 2007;45:S3-S11.
6. Methodology Committee of the Patient-Centered Outcomes Research Institute (PCORI). Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA*. 2012;307:1636-1640.
7. Critical Path Institute. Critical Path Institute project pipeline. Available at: http://www.c-path.org/project-pipeline.cfm. Accessed February 11, 2014.
8. UK National Health Service (NHS) Information Centre for Health and Social Care. Patient Reported Outcome Measures (PROMS). Available at: http://www.ic.nhs.uk/statistics-and-data-collections/hospital-care/patient-reported-outcome-measuresproms. Accessed February 11, 2014.
9. Aigner M, Forster-Streffleur S, Prause W, et al. What does the WHO-QOL-BREF measure? Measurement overlap between quality of life and depressive symptomatology in chronic somatoform pain disorder. *Soc Psychiatry Psychiatr Epidemiol*. 2006;41:81-86.
10. Berlim M, Fleck M. Quality of life and major depression. In: Ritsner MS, Awad AG, eds. *Quality of Life Impairment in Schizophrenia, Mood, and Anxiety Disorders*. Dordrecht, the Netherlands: Springer; 2007.
11. Moore M, Hofer S, McGee H, et al. Can the concept of depression and quality of life be integrated using a time perspective? *Health Qual Life Outcomes*. 2005;3:1-10.
12. Peselow ED, Karamians R, Lord M, et al. The discrepancy between patients and informants on clinician-rated measures in major depressive disorder: implications for clinical trials and clinical practice. *Int Clin Psychopharmacol*. 2014;29:111-115.
13. Rush AJ, Hiser W, Giles DE. A comparison of self-reported versus clinician-related symptoms in depression. *J Clin Psychiatry*. 1987; 48:246-248
14. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Soc Sci Med*. 1998;46:1569-1585.
15. IsHak WW, Greenberg JM, Balayan K, et al. Quality of life: the ultimate outcome measure of interventions in major depressive disorder. *Harv Rev Psychiatry*. 2011;19:229-239.
16. Pyne JM, Patterson TL, Kaplan RM, et al. Assessment of the quality of life of patients with major depression. *Psychiatr Serv*. 1997;48:224-230.
17. Ritsner M. The distress/protection vulnerability model of quality of life impairment syndrome. In: Ritsner MS, Awad AG, eds. *Quality of Life Impairment in Schizophrenia, Mood, and Anxiety Disorders*. Dordrecht, the Netherlands: Springer; 2007.
18. World Health Organization Quality of Life Group. Measuring quality of life, 1997. Available at: http://www.who.int/mental health/media/68.pdf. Accessed February 11, 2014.
19. Endicott J, Nee J, Harrison W, et al. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. *Psychopharmacol Bull*. 1993;29:321-326.
20. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). *Med Care*. 1992;30:473-483.
21. The WHOQOL Group. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Soc Sci Med*. 1998;46:1569-1585.
22. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199-208.

Remitters			Nonremitters				
IBI-D Post	IBI-D Change	*P*	n	IBI-D Pre	IBI-D Post	IBI-D Change	*P*
-2.40 (0.78)	-1.82 (1.12)	<0.0001	95	0.14 (1.03)	-0.20 (1.19)	-0.34 (1.02)	<0.0001
-2.50 (0.60)	-1.96 (0.82)	<0.0001	96	0.25 (0.87)	-0.16 (1.10)	-0.42 (0.97)	<0.0001
-2.22 (0.63)	-1.90 (0.91)	<0.0001	87	0.24 (0.93)	-0.26 (1.13)	-0.49 (0.86)	<0.0001
-2.68 (0.44)	-2.55 (1.07)	<0.0001	20	-0.11 (0.82)	-0.50 (1.12)	-0.39 (0.99)	<0.0001
-2.40 (0.64)	-1.64 (0.88)	<0.0001	95	-0.11 (0.95)	-0.44 (1.11)	-0.33 (0.91)	<0.0001
-2.30 (0.66)	-1.54 (0.90)	<0.0001	111	-0.15 (1.01)	-0.42 (1.29)	-0.27 (0.93)	<0.0001
-2.22 (0.49)	-1.90 (0.79)	<0.0001	37	-0.10 (0.89)	-0.49 (0.90)	-0.40 (1.01)	<0.0001

Table III. Continued

Resultados percibidos por el paciente antes y después del tratamiento del trastorno depresivo mayor

*Los resultados percibidos por el paciente (PROs) sobre la calidad de vida, el funcionamiento y la gravedad de los síntomas depresivos son importantes para la evaluación de la carga de enfermedad y para medir el impacto del tratamiento del trastorno depresivo mayor (TDM). Se intenta proporcionar un análisis detallado de los PROs antes y después del tratamiento del TDM a partir del gran estudio STAR*D (Sequenced Treatment Alternatives to Relieve Depression). Este análisis examina los PROs antes y después del tratamiento en el segundo nivel del STAR*D. Los datos completos sobre calidad de vida, funcionamiento y gravedad de los síntomas depresivos se analizaron para cada tratamiento del nivel 2 del STAR*D. Los PROs de calidad de vida, funcionamiento y gravedad de los síntomas depresivos mostraron un deterioro significativo después de fallar el ensayo con citalopram, un inhibidor selectivo de la recaptura de serotonina (en el nivel 1). Las siete opciones terapéuticas del nivel 2 tuvieron un impacto estadística (valores de p) y clínicamente (diferencias estandarizadas de Cohen [d de Cohen]) significativo en cuanto a calidad de vida, funcionamiento, gravedad de los síntomas depresivos y reducción en el cálculo de la carga de enfermedad. No hubo diferencias estadísticamente significativas entre las intervenciones. Sin embargo, una proporción significativa de pacientes mantuvo un deterioro en la calidad de vida y el funcionamiento después del tratamiento, y el efecto fue más pronunciado en aquellos que no remitieron. Los PROs son clave para la comprensión del impacto del TDM y para examinar los efectos de las intervenciones terapéuticas tanto en investigación como en clínica.*

Résultats rapportés par les patients avant et après traitement d'un épisode dépressif majeur

*Les résultats rapportés par les patients ou PRO (*Patient Reported Outcomes*) de qualité de vie (QdV), de fonctionnement, et de sévérité du symptôme dépressif sont importants dans l'évaluation du fardeau de l'épisode dépressif majeur (EDM) et de l'impact du traitement. Nous avons cherché à analyser de façon détaillée les PRO avant et après le traitement d'un EDM au cours de la deuxième étape de la grande étude STAR*D (*Sequenced Treatment Alternatives to Relieve Depression*). Les données complètes de QdV, fonctionnement et sévérité du symptôme dépressif sont analysées pour chaque traitement de l'étape 2 de STAR*D. Dans l'étape 1 de l'étude, après l'échec du citalopram, un inhibiteur sélectif de recapture de la sérotonine, les PRO de QdV, de fonctionnement et de sévérité du symptôme dépressif étaient très médiocres. L'impact des sept choix thérapeutiques de l'étape 2 sur la QdV, le fonctionnement, la sévérité du symptôme dépressif et la diminution du fardeau calculé de la maladie, a montré des différences statistiquement (valeurs de p) et cliniquement (différences standardisées de Cohen [*d de Cohen*]) positives. Il n'y a pas de différences statistiquement significatives entre les traitements. Une proportion importante de patients continue néanmoins à souffrir après le traitement, selon les résultats déclarés par les patients sur la QdV et le fonctionnement, et de façon plus prononcée chez ceux qui ne sont pas en rémission. Les PRO sont essentiels pour comprendre l'impact de l'EDM et pour observer les effets du traitement, à la fois pour la recherche et la pratique clinique.*

23. Cella D, Riley W, Stone A, et al; PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol*. 2010;63:1179-1194.

24. Mundt JC, Marks IM, Shear MK, Greist JH. The work and social adjustment scale: a simple measure of impairment in functioning. *Br J Psychiatry*. 2002;180:461-464.

25. Ustün TB, Chatterji S, Kostanjsek N, et al; WHO/NIH Joint Project. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull World Health Organ*. 2010;88:815-823.

26. Sheehan DV. *The Anxiety Disease*. New York, NY: Charles Scribner & Sons; 1983:151.

27. Endicott J, Nee J. Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacol Bull*. 1997;33:13-16.

28. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item quick inventory of depressive (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. 2003;54:585.

29. Beck AT, et al. *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation; 1996.

30. Zimmerman M, Chelminski I, McGlinchey JB, Posternak MA. A clinically useful depression outcome scale. *Compr Psychiatry*. 2008;49:131-140.

31. Radloff, LS. The CED-D scale: A self-report depression scale for research in the general population. *Appl Psychol Measures*. 1977;1:385-401.

32. Spitzer R, Kroenke K, Williams JB. Validation and utility of a self-report Version of PRIME-MD: the PHQ Primary Care Study. *JAMA*. 1999;282:1737-1744.

33. Teresi JA, Ocepek-Welikson K, Kleinman M, et al. Analysis of differential item functioning in the depression item bank from PROMIS: An item response theory approach. *Psychol Sci Q*. 2009;51:148-180.
34. IsHak WW, Greenberg JM, Saah T, et al. Development and validation of the Individual Burden of Illness Index for Major Depressive Disorder (IBI-D). *Adm Policy Ment Health*. 2013;40:76-86.
35. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *Am J Psychiatry*. 2006;163:1905-1917.
36. Thase ME, Friedman ES, Biggs MM, et al. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: a STAR*D report. *Am J Psychiatry*. 2007;164:739-752.
37. Rush AJ, Trivedi MH, Wisniewski SR, et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med*. 2006;354:1231-1242.
38. Trivedi MH, Fava M, Wisniewski SR, et al. Medication augmentation after the failure of SSRIs for depression. *N Engl J Med*. 2006;354:1243-1252.
39. 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;25:119-142.
40. 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;26:457-494.
41. Mundt JC, Marks IM, Shear MK, et al. The work and social adjustment scale: A simple measure of impairment in functioning. *Br J Psychiatry*. 2002;180:461-464.
42. Cohen RM, Greenberg J, IsHak WW. Incorporating multidimensional patient- reported outcomes of symptom severity, functioning, and quality of life in the Individual Burden of Illness index for Depression (IBI-D) to measure treatment impact and recovery in major depressive disorder. *JAMA Psychiatry*. 2013;70:343-350.
43. IsHak WW, Balayan K, Bresee C, et al. A descriptive analysis of quality of life using patient-reported measures in major depressive disorder in a naturalistic outpatient setting. *Qual Life Res*. 2013;22:585-596.
44. Lavori PW, Rush AJ, Wisniewski SR, et al. Strengthening clinical effectiveness trials: equipoise-stratified randomization. *Biol Psychiatry*. 2001;50:792-801.
45. Gaynes BN, Rush AJ, Trivedi MH, et al. The STAR*D study: treating depression in the real world. *Cleve Clin J Med*. 2008;75:57-66.
46. Rapaport MH, Clary C, Fayyad R, Endicott J. Quality-of-life impairment in depressive and anxiety disorders. *Am J Psychiatry*. 2005;162:1171-1178.
47. Schecter LE, Ring RH, Beyer CE, et al. Innovative approaches for the development of antidepressant drugs : current and future strategies. *NeuroRx*. 2005;2:590-611.
48. Fava GA, Savron G, Grandi S, et al. Cognitive-behavioral management of drug-resistant major depressive disorder. *J Clin Psychiatry*. 1997; 58:278-282.
49. de Mello MF, de Jesus Mari J, Bacaltchuk J, et al. A systematic review of research findings on the efficacy of interpersonal therapy for depressive disorders. *Eur Arch Psychiatry Clin Neurosci*. 2005;255:75-82.
50. Warden D, Rush AJ, Wisniewski SR, et al. What predicts attrition in second step medication treatments for depression?: a STAR*D Report. *Int J Neuropsychopharmacol*. 2009;12:459-473.
51. Warden D, Rush AJ, Wisniewski SR, et al. Income and attrition in the treatment of depression: a STAR*D report. *Depress Anxiety*. 2009;26:622-633.
52. Gaynes BN, Farley JF, Dusetzina SB, et al. Does the presence of accompanying symptom clusters differentiate the comparative effectiveness of second-line medication strategies for treating depression? *Depress Anxiety*. 2011;28:989-998.
53. IsHak WW, Greenberg JM, Cohen RM. Predicting relapse in major depressive disorder using patient-reported outcomes of depressive symptom severity, functioning, and quality of life in the individual burden of illness index for depression (IBI-D). *J Affect Disord*. 2013;151:59-65.