

Patient-Assessed Versus Physician-Assessed Disease Severity and Outcome in Patients With Nonspecific Pain Associated With Major Depressive Disorder

Koen Demyttenaere, M.D., Ph.D.; Durisala Desaiyah, Ph.D.;
Claude Petit, Ph.D.; Jens Croenlein, M.D.; and Stephan Brecht, M.D., Ph.D.

Objectives: This post hoc analysis compared how patients and physicians estimate disease severity and global improvement during 8 weeks of treatment for major depressive disorder (MDD) with associated nonspecific pain. In addition, predictors of pain and depression were identified.

Method: Data were derived from a double-blind, placebo-controlled, multicenter, European study (conducted from May 2005 to May 2006) in adult outpatients with MDD (DSM-IV criteria) and moderate pain not attributable to a diagnosed organic pain syndrome (Brief Pain Inventory-Short Form [BPI-SF] average pain score ≥ 3). Patients were randomly assigned to duloxetine 60 mg/day or placebo and treated for 8 weeks. Physicians were asked to rate severity of depression by using the Montgomery-Asberg Depression Rating Scale (MADRS) and the Clinical Global Impressions-Severity of Illness (CGI-S) and CGI-Improvement (CGI-I) scales. Patients were asked to assess pain using the BPI-SF, psychological symptomatology (9 domains including depression) with the Symptom Checklist-90-Revised (SCL-90-R), and overall improvement with the Patient Global Impression of Improvement (PGI-I). Multivariate linear regressions were performed as post hoc analyses to identify predictors of disease assessment at baseline and at the end of the study using a last-observation-carried-forward approach.

Results: All SCL-90-R domains improved during the 8 weeks of treatment. At baseline, the MADRS was associated only with the SCL-90-R obsessive-compulsive score, while the SCL-90-R depression score was associated with the BPI-SF average pain score and with many SCL-90-R subscores. The global impression of improvement was rated higher by the physicians than by the patients. At the end of the study, CGI-I was significantly associated with a decrease in depression severity (MADRS; $p < .0001$), younger age ($p = .0005$), and a decrease of the SCL-90-R interpersonal sensitivity score ($p = .0359$), but not with BPI-SF average pain. In contrast, patient-rated PGI-I was significantly associated with the SCL-90-R depressive domain ($p < .0001$), BPI-SF average pain ($p = .0003$), and the SCL-90-R anxiety domain ($p = .0041$) scores.

Conclusion: In patients with MDD associated with at least moderate nonspecific pain, physi-

cians consider mainly the change in depressive symptoms as measured by MADRS in their CGI-I ratings, while patients also consider pain, depression, and anxiety in their PGI-I ratings. When treating depression and assessing treatment outcome, a broad spectrum of symptoms needs to be monitored.

Trial Registration: clinicaltrials.gov
Identifier: NCT00191919

Prim Care Companion J Clin Psychiatry 2009;11(1):8-15

© Copyright 2009 Physicians Postgraduate Press, Inc.

Received May 15, 2008; accepted July 28, 2008. From the University Psychiatric Center, Catholic University of Leuven, Campus Gasthuisberg, Leuven, Belgium (Dr. Demyttenaere); Lilly Research Laboratories, Eli Lilly and Co., Indianapolis, Ind. (Dr. Desaiyah); Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, Conn. (Dr. Petit); Boehringer Ingelheim Pharma GmbH & Co., KG, Biberach, Germany (Dr. Croenlein); and Boehringer Ingelheim GmbH, Ingelheim, Germany (Dr. Brecht).

This study was sponsored by Boehringer Ingelheim GmbH, Ingelheim am Rhein, Germany, and Eli Lilly and Co., Indianapolis, Ind. The authors accept full responsibility for the conduct of this study, had full access to all data from this study, and participated in the decision to publish the data.

Presented at the 20th annual congress of the European College of Neuropsychopharmacology; Oct. 13-17, 2007; Vienna, Austria.

The authors wish to thank all the investigators and patients for their role in the conduct of this study.

Dr. Demyttenaere has served on the advisory boards of and received honoraria for services from Boehringer Ingelheim and Eli Lilly. Dr. Desaiyah is an employee and stock shareholder of Eli Lilly. Drs. Petit, Croenlein, and Brecht are employees of Boehringer Ingelheim.

Corresponding author and reprints: Koen Demyttenaere, M.D., Ph.D., University Psychiatric Center—Campus Gasthuisberg, Herestraat 49, 3000 Leuven, Belgium
(e-mail: Koen.Demyttenaere@med.kuleuven.be).

Many efficacy measures are used to assess change and outcome in antidepressant trials, and they differ in format and in content. The primary endpoint is usually defined as a change on a symptom scale reflecting the symptoms of the primary diagnosis.

A first distinction can be made between observer-rated instruments, such as the Hamilton Rating Scale for Depression (HAM-D) and the Montgomery-Asberg Depression Rating Scale (MADRS), and self-rated instruments, such as the Beck Depression Inventory (BDI) and the Symptom Checklist-90-Revised (SCL-90-R).

A second distinction can be made between instruments assessing syndromal symptomatology (HAM-D, MADRS, BDI, and SCL-90-R) and instruments assessing overall impression of clinical status and/or change, the latter generally being very short. The Clinical Global Impressions scale (CGI)¹ is such an “overall” assessment tool that generally shows good sensitivity to change, and its use has become so widespread that it is now a standard outcome measure in psychopharmacology trials. Since the term *global* gives the impression of comprehensiveness, this judgment may also be expected to be more representative for the actually observed and clinically relevant status or change.² But it has also been suggested that the global impression is an enigmatic and wavering term with poor psychometric validation and with semantic problems (for example, quality gaps with regard to intensity rating and no precise definitions of different categories) and logical problems (for example, a “moderately ill” patient cannot manifest “very much” improvement).³

The first CGI item assesses the clinician’s impression of the patient’s current illness severity (CGI-S) on a unipolar scale of 1 (not at all ill) to 7 (among the most extremely ill). The judgment is restricted within the range of the specific population under study.¹ The next 2 items of the CGI address the patient’s improvement from baseline, one rated by the clinician (Clinical Global Impressions-Improvement [CGI-I] scale) and the other by the patient (Patient Global Impression of Improvement [PGI-I]). Both items show a bipolar scaling from 1 (very much improved) to 7 (very much worse). The correlation between the PGI-I (patient global impression of change) and the CGI-I (clinical global impression of change) has been poorly investigated.

The term *global* does remain somewhat elusive and it is unclear which aspects of an individual’s functioning actually determine such ratings. Changes in symptom scores have been shown to account for 38% to 40% of the variance in CGI-S and 26% to 46% of the variance in CGI-I.⁴ The authors of this study suggest that other factors related to illness, such as appraisal of an individual’s distress, discomfort, and impairment, may account for the remaining variance in CGI ratings. In a study evaluating the CGI scale among individuals with social anxiety disorder, it was indeed shown that CGI-S and CGI-I were not only strongly correlated with observer- or self-rated severity of anxious symptoms but also with observer- or self-rated severity of comorbid depressive symptoms.⁵ The latter finding is important since it shows a limitation of using just 1 symptom scale for assessing treatment outcome. Indeed, the main symptom scale used usually reflects the symptoms of the main diagnosis and illustrates one of the weaknesses of the currently used diagnostic classification systems such as the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*: comorbid disorders are generally well assessed but

comorbid symptoms are not. Assessing outcome with only 1 symptom scale (reflecting the symptom cluster and severity of the primary “diagnosis”) could hence result in a “tunnel perspective” of the treating physician: once a physician makes a primary diagnosis of major depressive episode on the basis of the variety of symptoms a patient presents with, it is not well documented whether frequently comorbid symptoms (anxiety symptoms, obsessive-compulsive symptoms, interpersonal sensitivity, somatic nonpainful symptoms, and somatic painful symptoms) are further taken into account when assessing change during treatment.^{6,7} From this perspective, the HAM-D covers a broader set of symptoms than the MADRS, with the former scale also focusing on anxiety and somatic symptoms and the latter mainly focusing on depressive symptoms.⁸ The study by Zaider et al.⁵ suggests that more data are needed to understand what symptom changes contribute to the impression of global change and whether the same symptom changes predict the physician’s perspective (CGI-I) as well as the patient’s perspective (PGI-I).

This article aims to provide answers to these issues, including (1) the relation between different observer- and self-rated symptom scales (depressive and other comorbid psychological and somatic symptoms) in a population of outpatients with depression and pain, at baseline and during treatment; (2) the relation between global impression scales (CGI-I for the physician’s perspective and PGI-I for the patient’s perspective); and (3) the investigation of what symptom changes predict the CGI, the PGI, and the difference of both assessments.

METHOD

Study Design

The data for the current post hoc analyses were derived from a randomized, double-blind, placebo-controlled, parallel-group study conducted from May 2005 to May 2006 in 5 European countries (Belgium, Finland, France, Germany, and Slovakia) in outpatients with major depressive disorder (MDD) and at least moderate pain.⁹ The study duration was 10 weeks consisting of an 8-week treatment period and a 2-week tapering period. The treatments were placebo or duloxetine 60 mg/day, escalated from 30 mg/day after the first study week, followed by tapering after 7 weeks of treatment at 60 mg/day to 30 mg/day for 2 weeks. Eli Lilly and Company and Boehringer Ingelheim GmbH sponsored this study (ClinicalTrials.gov registration #NCT00191919).

Patients

A total of 321 adult male or female outpatients (duloxetine, N = 156; placebo, N = 165) with a diagnosis of MDD as defined by the DSM-IV¹⁰ were included. The diagnosis was confirmed by the Mini-International

Neuropsychiatric Interview (MINI, version 5.0.0), a standardized diagnostic interview based on DSM-IV criteria.¹¹ At baseline, all patients had a depression severity (MADRS) total score of ≥ 20 and at least a moderate pain (Brief Pain Inventory-Short Form [BPI-SF]) score of ≥ 3 for the “24-hour average pain” item.¹² At screening and at baseline, all patients had to be at least “moderately ill” as measured by a score of ≥ 4 on the CGI-S scale. The patients were devoid of any organic pain syndrome based on present and past medical history. Furthermore, patients must have had 1 previous depressive episode in their past medical history. Exclusion criteria and reasons for study exclusion have been discussed by Brecht et al.⁹ All patients signed an informed consent approved by the institutional ethical review board at each investigational site before their enrollment. The study was conducted in accordance with the Declaration of Helsinki.

Outcome Measures

Physician-rated scales. Depression severity was assessed with the MADRS scale¹³ at screening (visit 1) and at all visits. The scale includes 10 items that assess the core symptoms of depression. The items are scored from 0 to 6 with a maximum total score of 60. Reduction in score is a measure of symptom improvement.

The physicians used the CGI-S and CGI-I scales to evaluate the global disease severity and global improvement of patients.¹ The severity of the disease was assessed by the same evaluator throughout the study using the CGI-S item on a scale of 1 (normal) to 7 (most extremely ill) from screening to week 8; the CGI-I was used to assess the patient’s “global change” on a scale of 1 (very much improved) to 7 (very much worse) from week 1 through week 8. For the CGI-I and the PGI-I, the following categorizations were made: “improved” (scores 1 and 2), “stable” (scores 3, 4, and 5), or “worsened” (scores 6 and 7).

Patient-rated scales. The SCL-90-R was designed to characterize the global symptomatology and psychological distress of psychiatric outpatients^{14–16} and gives 9 domains, including somatization, obsessive-compulsive symptomatology, interpersonal sensitivity, depression, anxiety, anger/hostility, phobic anxiety, paranoid ideation, and psychoticism. The SCL-90-R was administered from baseline through week 8.

The BPI-SF¹² is a patient-rated instrument that measures pain intensity as well as other pain aspects on an 11-point Likert scale ranging from 0 (no pain or interference) to 10 (most severe pain or complete interference). Item 5 (average pain in the last 24 hours) of the BPI-SF was assessed at screening, baseline, and all other study visits. The BPI-SF interference questions were included to determine how much patients were limited by pain in their daily functioning (via 7 items: general activity,

mood, walking, normal work, relations with others, sleep, and enjoyment of life).

The patients self-rated their “global change,” ranging from 1 (very much better) to 7 (very much worse) using the PGI-I scale at weeks 3 through 8.

Statistical Analyses

All the analyses were conducted on the *full analysis set* (all patients who have received at least 1 dose of study medication and who have at least a baseline and a post-baseline value available for efficacy evaluation).

Cochran-Mantel-Haenszel tests adjusted on centers were performed to compare percentage of responders on CGI-I and PGI-I at endpoint (using a last-observation-carried-forward [LOCF] approach).

Multivariate linear regressions were performed as post hoc analyses to identify predictors of disease assessment at baseline and at the end of the study (LOCF). Models were fit to raw data with no stepwise approach.

The evolution of SCL-90-R variables between the treatment groups was assessed with a maximum likelihood-based, mixed-effects model repeated-measures analysis using all the observations at each postbaseline visit over 8 weeks of double-blind treatment. The model included the fixed categorical effects of treatment, center, visit, and treatment-by-visit interaction, as well as the continuous fixed covariates of baseline score and baseline score-by-visit interaction. The covariance structure to model the within-patient errors was unstructured; the Kenward-Roger method was used to estimate denominator degrees of freedom; and type III sum of squares for the least squares mean was used.

RESULTS

Relation Between Observer-Rated and Self-Rated Symptom Scales (depressive and other comorbid psychological and somatic symptoms)

Pain severity, interference of pain with functioning, and psychological variables from baseline to endpoint. The baseline and endpoint (LOCF) scores for all observer-rated and self-rated variables, including BPI-SF subscales, MADRS, and SCL-90-R domains, are presented in Table 1. Unadjusted means and standard error (SE) of the mean are displayed. The significant superior effect of duloxetine over placebo in reducing MADRS and BPI-SF scores has been reported previously.⁹

Observer-rated versus self-rated depression severity, SCL-90-R domains, and average pain severity/interference with functioning. Since the study sample were outpatients with MDD and moderate pain, the relationship between different symptom clusters was further investigated.

Table 2 shows the significant associations of the different SCL-90-R domains (excluding depressive symp-

Table 1. Baseline and Week 8 Scores for All Observer-Rated and Self-Rated Variables^a

Variable	Duloxetine		Placebo	
	Baseline	Week 8	Baseline	Week 8
BPI-SF average pain, mean (SE)	5.78 (0.13)	3.43 (0.21)	5.63 (0.12)	4.09 (0.19)
BPI-SF interference, mean (SE)	5.58 (0.15)	2.72 (0.20)	5.47 (0.14)	3.79 (0.19)
General activity	6.20 (0.17)	3.28 (0.23)	5.93 (0.15)	4.32 (0.22)
Mood	6.42 (0.17)	3.00 (0.23)	6.19 (0.16)	4.17 (0.22)
Walking ability	4.03 (0.23)	2.09 (0.22)	4.08 (0.22)	2.99 (0.22)
Work	5.69 (0.19)	2.85 (0.22)	5.53 (0.15)	3.97 (0.21)
Relations with people	4.88 (0.21)	2.11 (0.21)	4.78 (0.19)	3.11 (0.21)
Sleep	5.66 (2.20)	2.91 (0.24)	5.63 (0.19)	3.91 (0.23)
Enjoyment of life	6.18 (0.20)	2.79 (0.23)	6.14 (0.17)	4.08 (0.23)
MADRS, mean (SE)	29.91 (0.37)	14.78 (0.85)	29.30 (0.35)	19.06 (0.74)
SCL-90-R, mean (SE) (GSI)	1.46 (0.05)	0.81 (0.06)	1.51 (0.05)	1.08 (0.06)
Somatization	1.72 (0.76)	0.99 (0.82)	1.83 (0.70)	1.31 (0.83)
Obsessive-compulsive symptoms	1.80 (0.85)	1.03 (0.91)	1.83 (0.76)	1.34 (0.90)
Interpersonal sensitivity	1.35 (0.86)	0.71 (0.81)	1.41 (0.89)	0.97 (0.89)
Depressive symptoms	2.13 (0.76)	1.14 (0.99)	2.17 (0.72)	1.51 (0.91)
Anxiety symptoms	1.49 (0.85)	0.81 (0.84)	1.59 (0.78)	1.10 (0.89)
Anger/hostility	0.88 (0.82)	0.47 (0.66)	0.99 (0.93)	0.71 (0.89)
Phobic anxiety	0.83 (0.46)	0.46 (0.72)	0.82 (0.79)	0.63 (0.79)
Paranoid ideation	1.14 (0.96)	0.64 (0.82)	1.12 (0.90)	0.80 (0.88)
Psychoticism	0.84 (0.67)	0.48 (0.64)	0.87 (0.70)	0.66 (0.75)
CGI-S, n (%) ^b				
Moderate	130 (83.3)	78 (50.0)	134 (83.2)	104 (65.4)
Severe	26 (16.7)	6 (3.9)	27 (16.8)	11 (6.9)

^aUnadjusted means and standard errors (SE) of the mean at baseline and at visit 8 (using last-observation-carried-forward imputation method) are displayed in this table.

^bN = 156 duloxetine, N = 161 placebo (baseline); N = 156 duloxetine, N = 159 placebo (endpoint).

Abbreviations: BPI-SF = Brief Pain Inventory-Short Form, CGI-S = Clinical Global Impressions-Severity of Illness scale, GSI = general severity index, MADRS = Montgomery-Asberg Depression Rating Scale, SCL-90-R = Symptom Checklist-90-Revised.

Table 2. Regression Analysis Showing the Significant Predictors of the MADRS at Baseline and at Week 8^{a,b}

MADRS	BPI-SF Average Pain	SCL-90-R Obsessive-Compulsive	SCL-90-R Interpersonal Sensitivity	SCL-90-R Anxiety	SCL-90-R Phobic Anxiety	SCL-90-R Paranoid Ideation
Baseline	...	3.69
Week 8	4.06	3.92	2.57	2.64	-3.79	-2.67

^aThe t values from the linear multivariate regression model are displayed in this table; only t values that are significant at $p \leq .05$ are listed.

^bCovariates taken into account were the following: all the SCL-90-R subdomains except SCL-90-R depressive domain, number of pains, and BPI-SF average pain.

Abbreviations: BPI-SF = Brief Pain Inventory-Short Form, MADRS = Montgomery-Asberg Depression Rating Scale, SCL-90-R = Symptom Checklist-90-Revised).

Symbol: ... = not significant.

Table 3. Regression Analysis Showing the Significant Predictors of SCL-90-R Depression at Baseline and at Week 8^{a,b}

SCL-90-R Depression	BPI-SF Average Pain	SCL-90-R Obsessive-Compulsive	SCL-90-R Interpersonal Sensitivity	SCL-90-R Anxiety	SCL-90-R Phobic Anxiety	SCL-90-R Paranoid Ideation
Baseline	2.40	6.06	5.18	5.10	-3.05	...
Week 8	3.45	7.55	7.13	3.97	-4.09	-4.05

^aThe t values from the linear multivariate regression model are displayed in this table; only t values that are significant at $p \leq .05$ are listed.

^bCovariates taken into account were the following: all the SCL-90-R domains except SCL-90-R depressive domain, number of pains, and BPI-SF average pain.

Abbreviations: BPI-SF = Brief Pain Inventory-Short Form, SCL-90-R = Symptom Checklist-90-Revised,

Symbol: ... = not significant.

tomatology) with the observer-rated depression severity (MADRS) at baseline and at week 8 (linear-regression models). At baseline, only a higher SCL-90-R obsessive-compulsive symptom score was significantly associated with a higher baseline MADRS score. At week 8, the MADRS was significantly associated with BPI-SF average pain and with more psychopathological dimensions:

positively with obsessive-compulsive symptomatology, interpersonal sensitivity, and anxiety, and negatively with phobic anxiety and paranoid ideation.

Table 3 shows the significant associations of the different SCL-90-R domains (excluding depressive symptomatology) with the self-rated depression severity (SCL-90-R depressive domain) at baseline and at week 8. At baseline

Table 4. Significant (p values) Predictors of BPI-SF Average Pain and of BPI-SF Interference of Pain With Functioning at Baseline and at Week 8^a

BPI-SF	BPI-SF Average Pain	SCL-90-R Somatization	SCL-90-R Interpersonal Sensitivity	SCL-90-R Depression	SCL-90-R Anxiety	SCL-90-R Phobic Anxiety	SCL-90-R Paranoid Ideation	Age	Sex
Average pain at baseline ^b0018	.01580194
Average pain at week 8 ^c0010004	.0090008
Interference at baseline ^d	.0001	.0005	.042	.000103050056	...
Interference at week 8 ^d	.0001	.000100010368

^aOnly p values that are significant at ≤ .05 are listed.

^bCovariates taken into account at baseline were SCL-90-R domains 2 through 9, gender, and age.

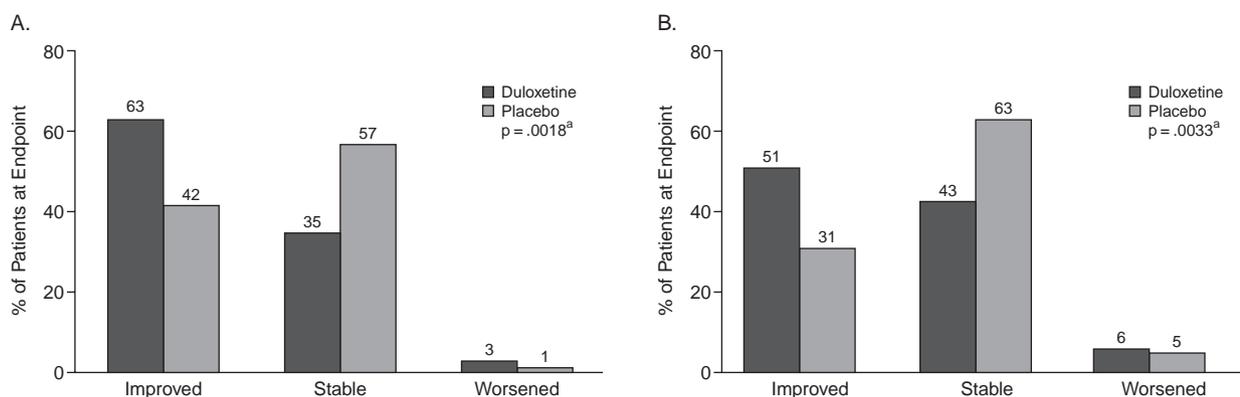
^cCovariates taken into account at week 8 were SCL-90-R domains 1 through 9, gender, and age.

^dBPI-SF interference is the sum of all interference sub-items (sum of the 7 items); covariates taken into account at baseline and week 8 were number of pains, BPI-SF average pain, SCL-90-R domains 1 through 9, gender, and age.

Abbreviations: BPI-SF = Brief Pain Inventory-Short Form, SCL-90-R = Symptom Checklist-90-Revised.

Symbol: ... = not significant.

Figure 1. Clinical Global Impression of Improvement (A) Versus Patient Global Impression of Improvement (B)



^aCochran-Mantel-Haenszel test stratified by study center (full analysis set, last observation carried forward).

as well as at week 8, a higher BPI-SF average pain score, a higher SCL-90-R obsessive-compulsive symptom score, a higher SCL-90-R interpersonal sensitivity score, a higher SCL-90-R anxiety score, and a lower SCL-90-R phobic anxiety score, as well as a lower SCL-90-R paranoid ideation score (but only at week 8), were significantly associated with the SCL-90-R depressive score.

Table 4 shows the significant associations of the different SCL-90-R domains with BPI-SF average pain and of BPI-SF interference (interference of pain with functioning) at baseline and at week 8. Higher baseline SCL-90-R depressive domain score, a lower baseline interpersonal sensitivity domain score, and a lower baseline phobic anxiety domain score were associated with a higher baseline average pain severity. A higher endpoint SCL-90-R depressive domain score, higher endpoint SCL-90-R somatization score, higher endpoint SCL-90-R anxiety score, and being female were associated with a higher average pain severity at endpoint.

A higher baseline average pain severity, higher baseline SCL-90-R somatization, lower baseline SCL-90-R interpersonal sensitivity, higher baseline SCL-90-R de-

pression, lower baseline SCL-90-R anxiety, and older age were associated with a higher interference of pain with functioning (average of the 7 interference of pain with functioning items). A higher endpoint pain severity, higher endpoint SCL-90-R somatization, higher endpoint SCL-90-R depression, and lower endpoint SCL-90-R paranoid symptomatology were associated with a higher endpoint interference of pain with functioning.

Clinical Global Impression of Improvement Versus Patient Global Impression of Improvement

The CGI-I scores showed that from the physician’s perspective, 63% of the patients on treatment with duloxetine and 42% of the patients taking placebo “improved.” On the other hand, the PGI-I scores showed that from the patient’s perspective, 51% of the patients on treatment with duloxetine and 31% of the patients taking placebo “improved” (Figures 1A and 1B).

There was a concordance between CGI-I and PGI-I in 161 patients (54%), but, in 113 patients (35%), the physician concluded a greater improvement than the patient, while in 32 patients (11%), the physician concluded a

Table 5. Regression Analyses of Symptom Clusters Predicting Improvement as Assessed by the Physicians Versus Patients at Week 8

Predictor of Improvement	t Values	p Values ^a
CGI-I (physician) predicted by ^b		
Decrease on MADRS	14.61	< .0001
Younger age	-3.51	= .0005
Decrease in SCL-90-R interpersonal sensitivity	2.11	= .0359
PGI-I (patient) predicted by ^c		
Decrease in SCL-90-R depression severity	6.22	< .0001
Decrease in BPI-SF average pain	3.66	= .0003
Decrease in SCL-90-R anxiety	2.89	= .0041

^aOnly p values that are significant at $\leq .05$ are listed.

^bCovariates included in the model were the change from baseline to week 8 in number of pains, BPI-SF average pain, MADRS total score, BPI-SF interference (all items), and SCL-90-R (all items); gender; and age.

^cThe decrease on MADRS was not considered as a covariate in this regression analysis. Covariates included in the model were the change from baseline to week 8 in number of pains, BPI-SF average pain, BPI-SF interference (all items), and SCL-90-R (all items); gender; and age.

Abbreviations: BPI-SF = Brief Pain Inventory-Short Form, CGI-I = Clinical Global Impressions-Improvement scale, MADRS = Montgomery-Asberg Depression Rating Scale, PGI-I = Patient Global Impression of Improvement, SCL-90-R = Symptom Checklist-90-Revised.

smaller improvement than the patient. The concordance was found in as many patients taking duloxetine as placebo (54% and 54%), but the discrepancy of improvement showed some numerical differences between the treatment groups. Physicians concluded there was a greater improvement than did the patients in 38% of patients taking duloxetine and in 33% of patients taking placebo. Physicians concluded a smaller improvement than the patients in 7% of patients taking duloxetine and in 13% of patients taking placebo.

Predictors of Global Assessment of Improvement

Regression analyses were performed in order to investigate what symptom clusters predicted "improvement" as assessed by the physicians versus the patients (Table 5). At study endpoint, a higher CGI-I was significantly predicted by a more pronounced decrease in depression severity (MADRS), younger age, and a decrease in SCL-90-R interpersonal sensitivity, but was not significantly predicted by BPI-SF average pain. In contrast, a higher PGI-I was predicted by a more pronounced decrease in SCL-90-R depression, a more pronounced decrease in BPI-SF average pain, and a more pronounced decrease in SCL-90-R anxiety.

Finally, a regression analysis was performed to determine whether the difference between CGI-I and PGI-I could be predicted by the sociodemographic, pain, and psychological variables (Table 6). A relatively higher impression of improvement by the physician than by the patient was predicted by a greater decrease in MADRS, a greater decrease in interference of pain with relations with people, and a greater decrease in SCL-90-R interpersonal

Table 6. Factors Predicting the Discrepancy^a of Patient-Assessed Disease Outcome (PGI-I) Versus Physician-Assessed Disease Outcome (CGI-I) at Week 8^b

Predictor of Discrepancy	t Values	p Values
Low discrepancy between CGI-I and PGI-I		
Low decrease in MADRS depression severity	-5.68	< .0001
Low decrease in BPI-SF pain interference with other people	-2.57	= .0106
Low decrease in SCL-90-R interpersonal sensitivity	-2.49	= .0135
High discrepancy between CGI-I and PGI-I		
Low decrease in BPI-SF average pain	2.37	= .0186
Low decrease in SCL-90-R patient-rated depression	3.55	= .0005

^aDiscrepancy is defined as PGI-I score minus CGI-I score at week 8.

^bCovariates included in the regression were the change from baseline to week 8 in number of pains, BPI-SF average pain, MADRS total score, BPI-SF interference (all items), and SCL-90-R (all items); gender; and age.

Abbreviations: BPI-SF = Brief Pain Inventory-Short Form, CGI-I = Clinical Global Impressions-Improvement scale, MADRS = Montgomery-Asberg Depression Rating Scale, PGI-I = Patient Global Impression of Improvement, SCL-90-R = Symptom Checklist-90-Revised.

sensitivity, as well as by a smaller decrease in SCL-90-R depression and a smaller decrease in BPI-SF average pain.

DISCUSSION

Patients with depression associated with moderate pain treated with duloxetine showed a greater improvement in BPI-SF average pain score and in MADRS score compared to patients treated with placebo.⁹ Interestingly, in this patient population, additional self-rated psychopathological domains assessed with the SCL-90-R improved over time. For obsessive-compulsive symptomatology, interpersonal sensitivity, depressive symptomatology, phobic anxiety, and paranoid ideation, the improvement was more pronounced in patients taking duloxetine than in patients taking placebo. Moreover, an interaction effect between duration of treatment and treatment group was found for obsessive-compulsive symptomatology, depressive symptomatology, somatization, and anxiety, which suggests that, with increasing duration of treatment, the improvement in duloxetine-treated patients became progressively more pronounced than in patients treated with placebo. This finding also suggests that, in pain-enriched MDD patients, a treatment (duloxetine or placebo) results in an improvement of several other psychopathological domains and that, for some of these domains, the effect is more pronounced with duloxetine than with placebo. It has been well documented that antidepressants are indeed also effective in anxiety disorders including obsessive-compulsive disorder. Duloxetine has been shown to be effective in subthreshold anxious symptoms in patients with MDD,^{17,18} but less has been reported on the other psychopathological domains

including interpersonal sensitivity, phobic anxiety, paranoid ideation, and somatization.

Two comments can be made on the basis of the present findings. First, the findings illustrate an often forgotten characteristic of the DSM classification system: in DSM-IV, there is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders or from no mental disorder. This outlook emphasizes the need to capture additional clinical information that goes beyond diagnosis, although dimensional systems communicate more clinical information because they report clinical attributes that might be subthreshold.¹⁰ Of course, a halo effect cannot be excluded since it could well be that depressed patients are so biased by their depressive mood that they tend to report negative feelings and symptoms.

Second, the data fit with the belief that antidepressants are broad spectrum antinervousness compounds since many are effective in comorbid conditions ranging from obsessive-compulsive disorders to premenstrual dysphoric disorder.¹⁹

The different regression analyses that were run in order to investigate the associations of the different SCL-90-R domains with observer-rated depression severity, self-rated depression severity, as well as with pain severity and pain interference, revealed clinically relevant information. Overall, the different self-rated psychopathological domains of the SCL-90-R were much more correlated with the self-rated depression severity than with the observer-rated depression severity. BPI-SF average pain, obsessive-compulsive symptomatology, interpersonal sensitivity, and anxiety scores all predicted higher self-rated depression severity, suggesting that they are all closely related. The more externally-oriented psychopathological domains (phobic anxiety and paranoid ideation) both predicted the lower internally oriented domain of self-rated depression severity. BPI-SF average pain severity and interference of pain with functioning also were more positively predicted by self-rated depression severity, anxiety, and the somatization subscale, while also being more negatively predicted by interpersonal sensitivity, phobic anxiety, and paranoid ideation. Additionally, female gender predicted a greater BPI-SF average pain severity and a more pronounced interference of pain with functioning.

The discrepancy between the physician's and the patient's global impressions of improvement was also investigated. While concordance was found in about half of the patients, overestimation of improvement by the physician or underestimation of improvement by the patient was found in 35% of the enrolled patients. That this discrepancy was not significantly different in patients treated with duloxetine compared with patients treated with placebo suggests that the discrepancy is not due to

the burden of adverse events (higher in duloxetine patients), as it is sometimes felt that the "clinical global impression" is a more global measure taking clinical improvement and tolerability into account. Indeed, it has been reported earlier that side effects only marginally affect the "clinical global impression."³

It is remarkable that, to the extent that societal and health economic imperatives invite a greater reliance on self-reported data, the discrepancies between observer- and self-rated depression severity have not been investigated more thoroughly. Underreporting of depression severity has been suggested in patients with less formal education and in older patients.²⁰ Overreporting of depression has been suggested in patients with higher scores on neuroticism and on oral and hysterical personality style; underreporting was found in patients with higher obsessiveness.^{21,22}

The present investigation sheds some new light on these discrepancies: when assessing global improvement, doctors seem to take into account mainly the decrease in observer-rated depression severity (MADRS), interpersonal sensitivity, and age, while patients take into account the change in depressive, anxious, and painful symptomatology. This result could well be in line with the finding that depressive disorders as well as anxiety disorders are closely related to chronic painful physical symptoms^{6,7} and that the presence of (painful) somatic symptoms can result in underrecognition and underdiagnosis of mental disorders since doctors do not systematically check the different symptom clusters but tend to rely on the most obvious one.²³ The data also suggest that physicians make a (DSM-IV) diagnosis on the basis of many symptoms a patient first presents; from there onwards, however, they mainly focus on the symptoms of the "main diagnosis" to assess improvement. Patients, on the other hand, think less in terms of "main diagnosis" and take into account the different concurrent symptom clusters (depressive, anxious, and [painful] somatic) when they have to assess improvement.

The limitations of this study include the fact that the data were calculated post hoc and not prespecified in the protocol of the study from which the data were derived. Physicians were asked to rate the diagnosis on the basis of the MADRS and not on the basis of other scales, suggesting the lack of a thorough diagnosis of depression. There also may be bias based on the fact that physicians were asked to follow the protocol per study visit only to avoid pseudo-behavioral-therapy-like effects with long and extensive study visits.

CONCLUSIONS

In patients with MDD associated with moderate pain, physicians consider mainly the change in depressive symptoms as measured by MADRS in their CGI-I assessment, while patients also consider pain, depression, and

anxiety in their PGI-I assessment. Making a main diagnosis of a MDD should not prevent physicians from taking into account often concurrent symptomatology such as anxious and painful physical symptoms.

Drug name: duloxetine (Cymbalta).

REFERENCES

- Guy W. ECDEU Assessment Manual for Psychopharmacology. US Dept Health, Education, and Welfare publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health; 1976:218-222
- Lehman AF, Babigian HM, Reed SK. The epidemiology of treatment for chronic and nonchronic mental disorders. *J Nerv Ment Dis* 1984; 172:658-666
- Beneke M, Rasmus W. "Clinical Global Impressions" (ECDEU): some critical comments. *Pharmacopsychiatry* 1991;25(4):171-176
- Leon AC, Shear MK, Klerman GL, et al. A comparison of symptom determinants of patient and clinician global ratings in patients with panic disorder and depression. *J Clin Psychopharmacol* 1993;13: 327-331
- Zaider TI, Heimberg RG, Fresco DM, et al. Evaluation of the Clinical Global Impressions scale among individuals with social anxiety disorder. *Psychol Med* 2003;33:611-622
- Demyttenaere K, Bonnewyn A, Bruffaerts R, et al. Comorbid painful physical symptoms and depression: prevalence, work loss, and help seeking. *J Affect Disord* 2006;92(2-3):185-193
- Demyttenaere K, Bonnewyn A, Bruffaerts R, et al. Clinical factors influencing the prescription of antidepressants and benzodiazepines: results from the European study of the epidemiology of mental disorders (ESEMeD) [published online ahead of print March 7, 2008]. *J Affect Disord* 2008;110(1-2):84-93. doi:10.1016/j.jad.2008.01.011
- Demyttenaere K, De Fruyt J. Getting what you ask for: on the selectivity of depression rating scales. *Psychother Psychosom* 2003;72(2): 61-70
- Brecht S, Courtecuisse C, Debieuvre C, et al. Efficacy and safety of duloxetine 60 mg once daily in the treatment of pain in patients with major depressive disorder and at least moderate pain of unknown etiology: a randomized controlled trial. *J Clin Psychiatry* 2007;68(11):1707-1716
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC: American Psychiatric Association; 1994
- Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59(suppl 20):22-33
- Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Singap* 1994;23:129-138
- Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382-389
- SCL-90 Self-Report Symptom Inventory. In: Guy W, ed. ECDEU Assessment Manual for Psychopharmacology, revised edition. US Dept Health, Education, and Welfare publication (ADM) 76-338. Rockville, Md: National Institute of Mental Health; 1976:314-330
- Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): a measure of primary symptom dimensions. *Mod Probl Pharmacopsychiatry* 1974;7:79-110
- Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974;19:1-15
- Hartford J, Kornstein S, Liebowitz M, et al. Duloxetine as an SNRI treatment for generalized anxiety disorder: results from a placebo- and active-controlled trial. *Int Clin Psychopharmacol* 2007;22:167-174
- Allgulander C, Hartford J, Russell J, et al. Pharmacotherapy of generalized anxiety disorder: results of duloxetine treatment from a pooled analysis of three clinical trials. *Curr Med Res Opin* 2007;23:1245-1252
- Leonard BE, Healy D. *Differential Effects of Antidepressants*. London, England: Martin Dunitz; 1999
- Enns MW, Larsen DK, Cox BJ. Discrepancies between self and observer ratings of depression: the relationship to demographic, clinical, and personality variables. *J Affect Disord* 2000;60(1):33-41
- Paykel ES, Prusoff BA. Relationships between personality dimensions: neuroticism and extraversion against obsessive, hysterical and oral personality. *Br J Soc Clin Psychol* 1973;12:309-318
- Duberstein PR, Heisel MJ. Personality traits and the reporting of affective disorder symptoms in depressed patients. *J Affect Disord* 2007;103: 165-171
- Tylee A, Walters P. Underrecognition of anxiety and mood disorders in primary care: why does the problem exist and what can be done? *J Clin Psychiatry* 2007;68(suppl 2):27-30