# Health-Related Quality-of-Life Measure Enhances Acute Treatment Response Prediction in Depressed Inpatients

Jeffrey M. Pyne, M.D.; Dale Bullock, B.S.; Robert M. Kaplan, Ph.D.; Thomas L. Smith, Ph.D.; J. Christian Gillin, M.D.; Shahrokh Golshan, Ph.D.; John R. Kelsoe, M.D.; and D. Keith Williams, Ph.D.



***Background:*** Many nonbiological variables are reported to predict treatment response for major depression; however, there is little agreement about which variables are most predictive.

***Method:*** Inpatient subjects (N = 59) diagnosed with current DSM-IV major depressive disorder completed weekly depressive symptom ratings with the Hamilton Rating Scale for Depression (HAM-D-17) and Beck Depression Inventory (BDI), and weekly health-related quality-of-life (HRQL) ratings with the Quality of Well-Being Scale (QWB). Acute responders were identified by a 50% decrease in HAM-D-17 score from baseline within 4 weeks of medication treatment. Predictor variables were initially chosen from a literature review and then tested for their association with acute treatment response.

***Results:*** An initial predictive model including age at first depression, admission BDI score, and melancholia predicted acute treatment response with 69% accuracy and was designated as the benchmark model. Adding the admission QWB index score to the benchmark model did not improve the prediction rate; however, adding the admission QWB subscales for physical and social activity to the benchmark model significantly improved acute treatment response prediction to 86% accuracy (p = .001).

***Conclusion:*** In addition to being designed for use in cost-effectiveness analyses, the QWB subscales appear to be useful HRQL variables for predicting acute inpatient depression treatment response.

*(J Clin Psychiatry 2001;62:261–268)*

*Received April 4, 2000; accepted Aug. 22, 2000. From the Department of Psychiatry, Central Arkansas Veterans Healthcare System and the Department of Psychiatry, University of Arkansas for Medical Sciences (Dr. Pyne); the Division of Biometry, University of Arkansas for Medical Sciences (Dr. Williams), Little Rock; and the Department of Family and Preventive Medicine (Dr. Kaplan) and the Department of Psychiatry (Mr. Bullock and Drs. Smith, Gillin, Golshan, and Kelsoe), University of California at San Diego, San Diego.*

*Supported by a Veterans Affairs Career Development Award and the VISN 16 MIRECC (Dr. Pyne), UCSD Mental Health Clinical Research Center grant MH 30914 (Dr. Gillin), UCSD Psychopharmacology and Psychobiology Fellowship grant MH 18399 (Michael Irwin, M.D.), and UCSD Clinical Research Center grant M01 RR00827 (Edward W. Holmes, M.D.).*

*A preliminary analysis of these data was presented at the Young Investigators poster session at the 151st annual meeting of the American Psychiatric Association; May 30–June 4, 1998; Toronto, Ontario, Canada.*

*The authors thank Marla Smith, R.N.; Kathy Resovsky, R.N.; and Debbie Greenfield, M.A., for assistance with data collection and Lesley Wetherell, B.S., for data management and statistical support.*

*Reprint requests to: Jeffrey M. Pyne, M.D., VAMC North Little Rock, Department of Psychiatry, 116F2/NLR, 2200 Fort Roots Dr., North Little Rock, AR 72114 (e-mail: pynejeffreym@exchange.uams.edu).*

Many nonbiological and biological factors have been reported to predict clinical response to antidepressant treatment, but there is little agreement on which factors are most predictive.[1] Examples of nonbiological predictors of treatment outcome include clinical,[2–8] sociodemographic,[9–11] and personality factors.[12–16] Examples of reported biological predictors include dexamethasone suppression,[17,18] sleep study tests,[19,20] thyrotropin stimulation,[21] dichotic listening,[22] and cerebrospinal fluid markers.[23]

Our review of the treatment response literature identified several clinical, sociodemographic, and personality factors as the most consistent predictors. The most consistent clinical factors appeared to be age at first depression,[11,24] severity of baseline depression,[25–27] melancholic symptoms,[2,28,29] acute onset of symptoms,[4,30] and comorbid Axis I or II disorders.[16,31–33] Socioeconomic status appeared to be the most consistent sociodemographic factor associated with treatment response.[7,34–36] Neuroticism appeared to be the most consistent personality factor.[37] However, neuroticism was not included in this study, because of the probable influence of acute illness on the measurement of this variable.[38]

The significance of discovering reliable treatment response predictors is highlighted by the economic and so-

cial costs of depressive disorders.[39,40] The use of reliable treatment predictors is important for both mental health treatment outcomes research[41,42] and efficient treatment planning. Accurate prediction of the timing of treatment response could assist clinicians in appropriately matching service intensity with likelihood of response, e.g., providing more frequent contact for those who do not respond acutely.

An understudied category of treatment response predictors is health-related quality of life (HRQL). Data from the Medical Outcomes Study showed that functional status as measured by Short Form-36 (SF-36) subscale scores was a strong predictor of treatment outcome for outpatients with depression.[43] These results led Wells et al.[43] to speculate that depressive symptoms and functional status may be stronger predictors of outcomes than the classification of depressive disorders. The SF-36 subscales have also been used to predict future health care utilization.[44] Criticisms of the SF-36, however, include its insensitivity as an HRQL measure in severely ill patient populations, resulting in a "floor effect"[45] that may limit its use as a predictor of inpatient depression treatment response.

The Quality of Well-Being scale (QWB)[46] (discussed in more detail in the Clinical Measures section below) is a general-purpose HRQL scale designed for use in cost-effectiveness analyses that provides a preference-weighted index score ranging from death (0.0) to perfect health (1.0). The QWB has been shown to be sensitive to a wide range of depression severity,[47] but has not previously been tested as an HRQL predictor of treatment response in depressed patients.

In this study, we constructed nonbiological predictor models for acute antidepressant response, with and without QWB index and subscale scores. We first constructed a "benchmark model" that included statistically significant nonbiological predictors other than the QWB scores. We hypothesized that (1) QWB index and subscale scores would be statistically significant stand-alone predictors of acute treatment response and that (2) the QWB index and subscale scores combined with the benchmark model would significantly improve acute treatment response prediction compared with the benchmark model alone.

## METHOD

### Subjects

A convenience sample of 66 subjects was enrolled from the University of California, San Diego Inpatient Mental Health Clinical Research Center (MHCRC) at the Veterans Affairs San Diego Healthcare System (San Diego, Calif.). Eligible veterans were included if they were between the ages of 20 and 70 years with a current DSM-IV diagnosis of a major depressive episode in the context of major depressive disorder (MDD) or bipolar I or II disorder following a formal structured diagnostic interview and confirmation at a weekly diagnostic consensus conference. The exclusion criteria included current diagnosis of substance abuse or dependence; current serious physical illness (e.g., unstable angina or seizure disorder); being unwilling or unable to provide a permanent address or the name, address, or phone number of at least one contact person; planning to leave the San Diego area within 1 year; and inability to read and complete self-administered questionnaires in English. Comorbid Axis I diagnoses other than current substance abuse or dependence were allowed in this study to better approximate a typical Veterans Affairs (VA) inpatient sample. Informed consent was obtained from all subjects prior to beginning the study.

### Design and Procedures

Weekly measures of depression severity and HRQL were collected using an observational study design. The initial ratings occurred within 2 days of admission and continued weekly for up to 4 weeks after medication treatment began. Baseline depression severity was defined as the mean Hamilton Rating Scale for Depression (HAM-D-17)[48,49] scores from admission (prehospital) and following the first week in the hospital to provide a more stable measurement of baseline depression severity. To account for the effects of being treated on a research unit (which may include a medication washout period), the 4-week medication treatment time period started after medication treatment was initiated. None of the subjects were part of a placebo-controlled study during the time of this study.

Acute treatment responders were predefined as achieving a 50% improvement in HAM-D score compared with their baseline HAM-D score (defined above) within 4 weeks of medication treatment.[50] This definition of acute treatment response was chosen because it is commonly used in depression treatment trials, it approximates the typical non–research unit procedure of discharging patients after documenting significant symptom improvement, and acute treatment response has been reported to predict later response.[51,52]

### Clinical Measures

All subjects were assessed with the Structured Clinical Interview for DSM-IV[53] after admission. Final Axis I through Axis V assessments were made during MHCRC diagnostic consensus conferences. The Axis II disorders assessed were borderline and antisocial personality disorder.

Weekly ratings included the HAM-D-17,[48,49] the Beck Depression Inventory (BDI),[54,55] and the QWB. The BDI was included as the depression symptom severity predictor because it is self-administered and therefore requires fewer clinical resources to administer than the interviewer-administered HAM-D-17. Using an intraclass correlation coefficient method,[56] the mean interrater reliability among

the HAM-D-17 interviewers was 0.92. The HAM-D-17 and QWB are both interviewer-administered measures and were administered by separate raters at each timepoint to preserve the independence of depression severity and HRQL assessment.

The QWB is an HRQL measure comprising 4 subscales: a symptom/problem complex (CPX) subscale and 3 functional subscales including physical activity (PAC), social activity (SAC), and mobility (MOB).[57,58] Each of the subscale scores is determined by preference weights derived from a representative community sample using a categorical rating scale method and a multiattribute utility model. The subscale scores are then subtracted from 1.0 (perfect health) to determine the QWB index score. The higher the subscale score, the greater the impairment associated with that subscale.

The QWB functional subscales are based on questions from national surveys including the Health Interview Survey, the Social Security Administration Survey of the Disabled, and a variety of other epidemiologic measures.[46] Items are organized into 3 subscales representing 3 distinct but related aspects of daily functioning. For example, MOB describes the ability to get around the community. Individuals who are most impaired are in a hospital for health reasons. Other levels of impairment represent the need for help in using public transportation and limitations in driving and travel. Those at the top level of functioning have no limitations in mobility. PAC relates to ambulation. The lowest level requires limitation to a wheelchair, bed, or chair. Intermediate levels represent trouble lifting, stooping, bending, or using stairs or use of a cane, crutches, or walker. The top level has no physical activity limitations. SAC describes limitation in activities of daily living. The most severe level requires help with self-care activities. Intermediate levels suggest limitations in major role performance or limitations in recreational or leisure activities. The top level has no social activity limitations.

The QWB index score is on a continuum between 0.0 and 1.0 representing death and perfect health, respectively. With 2 or more data points extending over a year, the QWB output can be converted to quality-adjusted life-year (QALY) units, which are the recommended units of effectiveness in cost-effectiveness analysis.[59] The QWB has been used extensively to assess the health status of patients with physical disorders both at baseline and longitudinally.[60] Examples of mean QWB scores for community controls and patients with physical health disorders include 0.81 for community controls,[57] 0.66 for patients with chronic obstructive pulmonary disease,[61] 0.60 for patients with rheumatoid arthritis on treatment with placebo,[62] and 0.46 for patients with major trauma at discharge.[63] The QWB has also been shown to be sensitive to cross-sectional and longitudinal differences in depression severity.[47,64]

### Nonbiological Predictors of Response to Acute Treatment

As outlined in the introduction above, clinical and sociodemographic acute treatment response predictors that were identified on the basis of a literature review were tested. The clinical predictors included age at first depression, presence of DSM-IV MDD with melancholic features, depression severity, depression chronicity, and the presence of comorbid Axis I and II disorders. The sociodemographic predictor tested was socioeconomic status, defined by level of education, work status, and score on the Hollingshead Two-Factor Index of Social Position.[36] The Hollingshead score is a 2-factor index score that combines level of formal education and level of occupation status into a single score. Lower scores indicate higher socioeconomic status. The HRQL variables tested were the QWB index score and subscale scores described above.

### Statistical Analysis

All data were analyzed using NCSS 2000 statistical software.[65] Continuous data approximated a normal distribution, and comparisons between acute responders and nonresponders were made using 2-tailed t tests. Categorical data were compared using the chi-square statistic. The association of acute treatment response and baseline predictor variables was evaluated using the Spearman correlation coefficient with 2-tailed significance tests, because treatment response was coded as a dichotomous variable. Logistic regression analyses were used to predict acute treatment response from baseline predictor data. Logistic regression analyses allowed the use of a common measurement of treatment response (50% decrease in depression severity) and a simultaneous examination of the effects of categorical and continuous independent variables (sociodemographic and clinical predictors) on a dichotomous outcome (acute treatment response and nonresponse).[66] For the final predictive model, sensitivity, specificity, and positive and negative predictive value are presented. This approach was chosen because there is not a true "gold standard" for acute treatment response comparison.

## RESULTS

Sixty-six subjects were enrolled in the study. Seven subjects were dropped from the analysis because they completed only the baseline ratings (prehospital and first week in hospital) and therefore could not be designated as acute responders or nonresponders: 4 were discharged against medical advice, and 3 had regular discharges but were lost to follow-up. There were no statistically significant differences between noncompleters and completers according to education, age at first depression, or prehospital HAM-D-17 and QWB scores. However, non-

**Table 1. Responder Versus Nonresponder Socioeconomic Comparison**

| | Responders (N = 31) | | Nonresponders (N = 28) | | |
|---|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD | p Value |
| Education, y | 13.5 | 2.3 | 13.5 | 2.0 | NS |
| Hollingshead score[a] | 4.6 | 1.1 | 4.9 | 0.7 | NS |
| Work status | | | | | |
| Employed/retired | 9 | 29.0 | 5 | 17.9 | |
| Unemployed | 22 | 71.0 | 23 | 82.1 | NS |

[a]Hollingshead Two-Factor Index of Social Position.

**Table 2. Responder Versus Nonresponder Prehospital Clinical Comparison[a]**

| | Responders (N = 31) | | Nonresponders (N = 28) | | |
|---|---|---|---|---|---|
| Variable | N | % | N | % | p Value |
| Melancholia | | | | | |
| Yes | 21 | 67.7 | 9 | 32.1 | |
| No | 10 | 32.3 | 19 | 67.9 | .006 |
| Chronic depression (2 y or more) | | | | | |
| Yes | 6 | 19.4 | 7 | 25.0 | |
| No | 25 | 80.6 | 21 | 75.0 | NS |
| Comorbid Axis I disorder | | | | | |
| Yes | 23 | 74.2 | 24 | 85.7 | |
| No | 8 | 25.8 | 4 | 14.3 | NS |
| Comorbid Axis II disorder | | | | | |
| Yes | 8 | 25.8 | 7 | 25.0 | |
| No | 23 | 74.2 | 21 | 75.0 | NS |
| | Mean | SD | Mean | SD | |
| Age at first depression | 38.1 | 16.0 | 29.0 | 12.2 | .02 |
| BDI score | 29.1 | 8.0 | 34.5 | 6.9 | .009 |
| HAM-D-17 score | 23.1 | 4.6 | 22.4 | 5.9 | NS |
| QWB | | | | | |
| Index score | 0.538 | 0.06 | 0.507 | 0.06 | NS |
| Symptom/problem complex score | 0.308 | 0.047 | 0.326 | 0.019 | NS |
| Mobility score | 0.016 | 0.025 | 0.028 | 0.031 | NS |
| Physical activity score | 0.057 | 0.025 | 0.047 | 0.034 | NS |
| Social activity score | 0.060 | 0.017 | 0.070 | 0.014 | NS |

[a]Abbreviations: BDI = Beck Depression Inventory, HAM-D-17 = 17-item Hamilton Rating Scale for Depression, QWB = Quality of Well-Being Scale.

completers had a lower mean ± SD prehospital BDI score (23.7 ± 5.6) than did study completers (31.6 ± 7.9; t = 3.4, df = 63, p = .008). The 59 subjects who completed the acute response assessments were included in the subsequent analyses. The mean age for the completers was 47.2 ± 10.4 years (range, 25–66 years); 85% (50/59) were men, and 83% (49/59) were white.

The following were the primary antidepressant medications used for acute treatment: 47% of subjects (28/59) took selective serotonin reuptake inhibitors; 17% (10/59), mood stabilizers; 12% (7/59), bupropion; 7% (4/59), venlafaxine; 7% (4/59), monoamine oxidase inhibitors; 5% (3/59), nefazodone; 3% (2/59), tricyclic antidepressants; 2% (1/59), electroconvulsive therapy. Of the subjects who completed the acute response assessments, 12% (7/59) were also enrolled in a medication treatment study in which patients were randomly assigned to bupropion or sertraline. The remaining 88% of subjects (52/59) were prescribed treatment solely at the discretion of the inpatient treatment team.

We recruited subjects with a diagnosis of a major depressive episode in the context of either MDD or bipolar I or II disorder because we were predicting treatment response over a short period of time and there is very little evidence in the literature regarding acute depression response predictors specific to bipolar disorder. In fact, we found no statistical difference in the proportion of MDD versus bipolar acute responders: 57% (25/44) versus 40% (6/15), respectively ($\chi^2 = 1.3$, df = 1, p = .26). Therefore, MDD and bipolar subjects were combined for analyzing the prediction models. In addition, there were no statistically significant differences between MDD and bipolar subjects' socioeconomic and clinical predictors described above, except that bipolar subjects experienced their first depressive episode at a younger mean age than MDD subjects (26.4 ± 14.0 vs. 36.3 ± 14.5; t = 2.3, df = 57, p = .03).

As shown in Table 1, acute responders and nonresponders did not differ significantly according to the socioeconomic status factors tested (education, work status, and Hollingshead score[36]).

The comparison of prehospital clinical variables by acute treatment response group is shown in Table 2. Prehospital depression severity and HRQL measurements were used instead of the mean of prehospital and first-week-in-the-hospital scores, because these potential predictor variables would allow the clinician to make acute response predictions early in the admission. Responders had significantly lower prehospital BDI scores than nonresponders (t = –2.72, df = 56, p = .009). In addition, responders were more likely to have melancholic features ($\chi^2 = 7.5$, df = 1, p = .006) and to have experienced their first depressive episode at an older age than nonresponders (F = 5.92, df = 58, p = .02). Univariate comparisons between responders and nonresponders according to presence of comorbid Axis I and II disorders, chronic MDD (2 years or more of meeting criteria for MDD), prehospital QWB index score, and QWB subscale scores were not statistically significant.

The Spearman correlation coefficients between statistically significant univariate prehospital predictors of acute treatment response are presented in Table 3. Older age at first depression, the presence of melancholic symptoms, and lower prehospital BDI scores were significantly associated with acute treatment response. In addition, there appeared to be no evidence of colinearity among the statistically significant clinical prehospital predictors based on the lack of correlation among the predictor variables themselves.

Logistic regression analyses were conducted using the statistically significant predictor variables described above. We first tested a model that consisted of age at first

**Table 3. Correlation Matrix for Significant Prehospital Predictors of Acute Treatment Response[a]**

| Variable | Age at First Depression | Melancholia | BDI Score |
|---|---|---|---|
| Response (no response = 0, response = 1) | 0.29* | 0.36** | –0.38** |
| Age at first depression | 1.0 | –0.02 | –0.19 |
| Melancholia (0 = not melancholic, 1 = melancholic) | ... | 1.0 | –0.18 |

[a]Abbreviation: BDI = Beck Depression Inventory. Spearman correlation coefficient with 2-tailed significance used.
*p < .05.
**p < .01.

**Table 4. Benchmark Model Predicting Acute Treatment Response[a]**

| | Observed | | |
|---|---|---|---|
| Predicted | Nonresponder (N = 27)[b] | Responder (N = 31) | |
| Nonresponder | 18 | 9 | |
| Column (%) | (66.7) | (29.0) | |
| Row (%) | (66.7) | (33.3) | |
| Responder | 9 | 22 | |
| Column (%) | (33.3) | (71.0) | |
| Row (%) | (29.0) | (71.0) | |
| Independent Prehospital Variable | β[c] | Wald[d] | p |
| Age at first depression | 0.05 | 5.1 | .02 |
| Melancholia (0 = not melancholic, 1 = melancholic) | 1.60 | 6.0 | .01 |
| BDI score | –0.08 | 3.6 | .06 |

[a]Abbreviation: BDI = Beck Depression Inventory.
Model $\chi^2 = 18.14$, df = 3, p = .0004; model $r^2 = 25\%$; sensitivity for detecting nonresponders = 67%; specificity for detecting responders = 71%; positive predictive value = 67%; negative predictive value = 71%.
[b]One nonresponding subject is not included because no prehospital BDI score was available.
[c]β = logistic-model regression coefficient.
[d]Wald = a chi-square test obtained by $[\beta/SE(\beta)]^2$; see Hosmer and Lemeshow.[67]

depression, melancholia, and prehospital BDI score. This model resulted in a significant equation (model $\chi^2 = 18.14$, df = 3, p = .0004) and predicted acute response with 69% accuracy. Age at first depression and melancholia were statistically significant predictors (p = .02 and p = .01, respectively), and BDI was a marginally statistically significant predictor (p = .06). We concluded that this 3-factor model would be the benchmark model for future comparisons (Table 4).

We next tested the QWB index score and QWB subscale scores on their own in separate models. The predictive model with the QWB index score alone was not statistically significant (model $\chi^2 = 2.62$, df = 1, p = .11). The model with the QWB subscale scores alone was statistically significant (model $\chi^2 = 13.03$, df = 4, p = .01), predicted acute treatment response with 71% accuracy, and the SAC and PAC subscales were statistically significant predictors (p = .04 and p = .03, respectively).

**Table 5. Optimal Model Predicting Acute Treatment Response[a]**

| | Observed | | |
|---|---|---|---|
| Predicted | Nonresponder (N = 27)[b] | Responder (N = 31) | |
| Nonresponder | 24 | 5 | |
| Column (%) | (88.9) | (16.1) | |
| Row (%) | (82.8) | (17.2) | |
| Responder | 3 | 26 | |
| Column (%) | (11.1) | (83.9) | |
| Row (%) | (10.3) | (89.7) | |
| Independent Prehospital Variable | β[c] | Wald[d] | p |
| Age at first depression | 0.07 | 6.2 | .01 |
| Melancholia | 2.7 | 8.0 | .005 |
| BDI score | –0.11 | 4.1 | .04 |
| Social activity score | –50.4 | 5.7 | .02 |
| Physical activity score | 44.3 | 8.7 | .003 |

[a]Abbreviation: BDI = Beck Depression Inventory. Model $\chi^2 = 31.89$, df = 5, p < .000006; model $r^2 = 38\%$; sensitivity for detecting nonresponders = 89%; specificity for detecting responders = 84%; positive predictive value = 83%; negative predictive value = 90%.
[b]One nonresponding subject is not included because no prehospital BDI score was available.
[c]β = logistic-model regression coefficient.
[d]Wald = a chi-square test obtained by $[\beta/SE(\beta)]^2$; see Hosmer and Lemeshow.[67]

In the next step, we added the QWB scores to the benchmark model. We first added the QWB index score to the benchmark model, and this resulted in a significant equation (model $\chi^2 = 18.15$, df = 4, p = .001) that was not statistically different from the benchmark model (age at first depression, BDI score, and melancholia) (model $\chi^2$ difference = .01, df = 1, p = .92). The QWB index score was not a significant predictor in this model (p = .91).

Adding the statistically significant QWB subscales, SAC and PAC, to the benchmark model also resulted in a significant equation (model $\chi^2 = 31.89$, df = 5, p = .000006). The addition of the QWB subscales resulted in a statistically more robust prediction model than did the benchmark model (model $\chi^2$ difference = 13.75, df = 2, p = .001) and predicted acute treatment response with 86% accuracy (Table 5). The combination of the clinical and QWB subscale variables thus resulted in an overall 17% improvement in prediction of acute treatment response (25% improvement in the prediction rate) and a 22% improvement in prediction of nonresponse (33% improvement in the prediction rate).

## DISCUSSION

Predictors of relapse are an important area of mental health outcomes research.[42,68] Wells et al.[43] showed that depressive symptom severity and HRQL were powerful predictors of treatment outcome for depression at 1- and 2-year follow-up. Lyness et al.[69] noted that symptom and functional measures tap related but differentiable outcome constructs and therefore may independently predict treatment response.

In this study, we found the direction of the relationship between age at first depression and melancholia with acute treatment response to be consistent with the literature. The association we found between depression severity and acute treatment response was mixed. The prehospital HAM-D-17 score was not associated with acute treatment response (see Table 2). The prehospital BDI score had a statistically significant univariate relationship with acute response consistent with the literature and was marginally statistically significant in the benchmark multivariate prediction model. In addition, we found the SAC and PAC subscale scores of a generic HRQL instrument, the QWB, to be among the strongest predictors of acute response to inpatient antidepressant treatment when compared with commonly cited nonbiological treatment response predictors. Specifically, lower SAC scores (indicating less social impairment) predicted acute response, and higher PAC scores (greater physical impairment) predicted acute response.

The hypotheses that the QWB index score would be a significant predictor of acute treatment response both alone and when added to the benchmark model were not supported. A possible reason that the QWB index score was not a strong predictor of acute treatment response is that the QWB index score is determined by 4 independent subscale scores. Therefore, the subscales may dilute or cancel out the ability of the other subscales to predict acute treatment response. In fact, this explanation is quite likely, because the CPX and MOB subscales were not significant predictors, but the SAC and PAC subscales were significant predictors, and their regression coefficients were of opposite sign. Because the QWB index score is calculated by subtracting the subscale scores from 1.0, the effect of subscales that predict response in opposite directions could decrease the predictive power of the QWB index score.

The hypotheses that the individual QWB subscales would be significant predictors of acute treatment response both alone and when added to the benchmark model were supported. Specifically, the addition of the QWB subscales SAC and PAC resulted in a 17% improvement in prediction overall (25% improvement in the prediction rate) and a 22% improvement in prediction of nonresponders (33% improvement in the prediction rate). One reason why the subscale scores are stronger predictors than the QWB index score could be that subscale scores contain more content-specific items and therefore may target more specific domains of acute treatment response prediction than an overall index score. Also, individual subscale scores are not directly affected by the direction of the predictive effect of other subscales.

The opposing direction of acute treatment response prediction for SAC and PAC subscales is intriguing. According to the QWB questionnaire, a higher score on the SAC subscale means greater impairment in occupational, leisure, and self-care functioning. According to the data presented here, greater impairment on the SAC subscale predicts nonresponse. There is evidence in the literature for a relationship between depression severity as measured by the 21-item HAM-D and social functioning as measured by the Social Adjustment Scale.[70] However, we found a nonsignificant correlation between SAC and HAM-D-17 scores ($r = .11$, $p = .4$) and a marginally statistically significant correlation between SAC and BDI scores ($r = .24$, $p = .08$). In addition, the SAC subscale remains a significant predictor of acute treatment response when added to the benchmark model that includes the BDI. The SAC subscale may therefore provide additional predictive information that is not found in the depression severity ratings.

A higher score on the PAC subscale means greater physical activity impairment (for example, spending most of the day in a bed, chair, or couch for health reasons; or impaired physical movement) and predicts treatment response. The relationship between greater physical activity impairment and acute treatment response could be explained if PAC was correlated with another predictor. The only measured predictor included in the models that was correlated with PAC was the SAC subscale ($r = .37$, $p = .005$). Another possible explanation for the relationship is that if a high PAC score is associated with a low-energy and low-motivation state, it may decrease with the structured inpatient (milieu and medication) treatment environment. This explanation is partially supported by the significant correlation between prehospital HAM-D-17 (measuring predominantly neurovegetative symptoms of depression) and PAC scores ($r = .27$, $p = .04$) and the nonsignificant correlation with prehospital BDI score (measuring an array of cognitive, behavioral, and somatic symptoms of depression) and PAC scores ($r = .18$, $p = .17$). Of course, another explanation could be that PAC is a proxy for an as yet unmeasured predictor.

The sensitivity and specificity of the final 5-factor model for detecting nonresponse were 89% and 84%, respectively (see Table 5). Among community subjects diagnosed with major depression in the Epidemiologic Catchment Area study, the sensitivity and specificity to predict diagnostic status 1 year later were 17% and 97%, respectively, using a large number of sociodemographic and clinical variables.[6] The reason for mentioning this study is not for direct comparison, but to highlight a conclusion made by Sargeant et al.[6] about the need to look beyond the traditional variables that are used to predict clinical course.

From the perspective of designing treatment plans, predicting which patients will be nonresponders could be useful. For example, the clinician could plan to invest more resources (e.g., more frequent follow-up, patient and family education, or medication or psychotherapy augmentation) to achieve treatment response for patients predicted to be nonresponders. In addition, the clinician

could more accurately inform the treatment expectations of the predicted nonresponder patients and their family members or caregivers.

In this sample, the ability to detect acute treatment nonresponse increased from 67% to 89% with the addition of the QWB subscales to the benchmark model. Because the predictive value of any model depends on the prevalence of the condition (in this case, the prevalence of acute treatment nonresponse), the utility of the prediction model presented here depends on the acute response rate for a given clinical setting. For example, based on the results presented here, if the acute response rate is 50% (close to that found in our sample), the positive predictive value for predicting nonresponse would be 85%. If the response rate were decreased to 25%, the positive predictive value for predicting nonresponse would increase to 94%. If the response rate were increased to 75%, the positive predictive value for predicting nonresponse would decrease to 65%. Therefore, the model presented here for detecting treatment nonresponse appears to be best suited for clinical settings where the acute treatment response rate is approximately 50% or less.

There are important limitations to this study. For example, our sample size is relatively small and includes MDD and bipolar subjects. Given this limitation however, a model is emerging for predicting acute inpatient depression treatment response with QWB subscale scores combined with easily obtained clinical data. In addition, using the final 5-factor model with only those subjects diagnosed with MDD (N = 44), the model predicts with 81% accuracy overall. A second limitation is that the sample was drawn from an inpatient VA setting and includes severely disabled patients, which may limit the generalizability of the findings to less disabled patient populations. A third limitation is that we cannot draw conclusions about which elements of treatment result in acute treatment response. Future studies using randomized, controlled treatment designs could address this question.[50]

The QWB scale was designed for use as a generic method for determining cost-effectiveness ratios of health care interventions across the physical health/mental health spectrum. In addition to their use in cost-effectiveness analyses, our study suggests that QWB subscale scores may also be valuable in predicting major depressive episode acute inpatient treatment response in major depressive disorder. Further investigation of the use of the QWB subscale scores as treatment response predictors is warranted.

*Drug names:* bupropion (Wellbutrin), nefazodone (Serzone), sertraline (Zoloft), venlafaxine (Effexor).



## REFERENCES

1. Dunner DL. Comment and review. In: Goodnick PJ, ed. Predictors of Treatment Response in Mood Disorders. Washington, DC: American Psychiatric Press; 1996:199–214
2. Heiligenstein JH, Tollefson GD, Faries DE. Response patterns of depressed outpatients with and without melancholia: a double-blind, placebo-controlled trial of fluoxetine versus placebo. J Affect Disord 1994;30:163–173
3. Taylor S, McLean P. Outcome profiles in the treatment of unipolar depression. Behav Res Ther 1993;31:325–330
4. Keller MB, Lavori PW, Mueller TI, et al. Time to recovery, chronicity, and levels of psychopathology in major depression: a 5-year prospective follow-up of 431 subjects. Arch Gen Psychiatry 1992;49:809–816
5. Sotsky SM, Glass DR, Shea MT, et al. Patient predictors of response to psychotherapy and pharmacotherapy. Am J Psychiatry 1991;148: 997–1008
6. Sargeant JK, Bruce ML, Florio LP, et al. Factors associated with 1-year outcome of major depression in the community. Arch Gen Psychiatry 1990;47:519–526
7. Keller MB, Klerman GL, Lavori PW, et al. Long-term outcome of episodes of major depression: clinical and public health significance. JAMA 1984;252:788–792
8. Keller MB, Shapiro RW. Major depressive disorder: initial results from a one-year prospective naturalistic follow-up study. J Nerv Ment Dis 1981; 169:761–767
9. Friedman RA, Parides M, Baff R, et al. Predictors of response to desipramine in dysthymia. J Clin Psychopharmacol 1995;15:280–283
10. Zubenko GS, Mulsant BH, Rifai AH, et al. Impact of acute psychiatric inpatient treatment on major depression in late life and prediction of response. Am J Psychiatry 1994;151:987–994
11. Veiel HOF, Kuhner C, Brill G, et al. Psychosocial correlates of clinical depression after psychiatric in-patient treatment: methodological issues and baseline differences between recovered and non-recovered patients. Psychosom Med 1992;22:415–427
12. Nelson E, Cloninger CR. Exploring the TPQ as a possible predictor of antidepressant response to nefazodone in a large multi-site study. J Affect Disord 1997;44:197–200
13. Katon W, Von Korff M, Lin E, et al. Methodologic issues in randomized trials of liaison psychiatry in primary care. Psychosom Med 1994;56: 97–103
14. Hoencamp E, Haffmans PM, Duivenvoorden H, et al. Predictors of (non-) response in depressed outpatients treated with a three-phase sequential medication strategy. J Affect Disord 1994;31:235–246
15. Scott J, Eccleston D, Boys R. Can we predict the persistence of depression? Br J Psychiatry 1992;161:633–637
16. Eccleston D, Scott J. Treatment, prediction of relapse and prognosis of chronic primary major depression. Int Clin Psychopharmacol 1991; 6(suppl 2):3–10
17. Schweitzer I, Maguire KP, Gee AH, et al. Prediction of outcome in depressed patients by weekly monitoring with the dexamethasone suppression test. Br J Psychiatry 1987;151:780–784
18. Ribeiro SCM, Tandon R, Grunhaus L, et al. The DST as a predictor of outcome in depression: a meta-analysis. Am J Psychiatry 1993;150: 1618–1629
19. Heiligenstein JH, Faries DE, Rush AJ, et al. Latency to rapid eye movement sleep as a predictor of treatment response to fluoxetine and placebo in nonpsychotic depressed outpatients. Psychiatry Res 1994;52:327–339
20. Simons AD, Thase ME. Biological markers, treatment outcome, and 1-year follow-up in endogenous depression: electroencephalographic sleep studies and response to cognitive therapy. J Consult Clin Psychol 1992;60:392–401
21. Langer G, Koinig G, Hatzinger R, et al. Response of thyrotropin to thyrotropin-releasing hormone as predictor of treatment outcome: prediction of recovery and relapse in treatment with antidepressants and neuroleptics. Arch Gen Psychiatry 1986;43:861–868
22. Bruder GE, Otto MW, McGrath PJ, et al. Dichotic listening before and after fluoxetine treatment for major depression: relations of laterality to therapeutic response. Neuropsychopharmacology 1996;15:171–179
23. Bunney WE, Garland-Bunney B, Patel SB. Biological markers in depression. Psychopathology 1986;19(suppl 2):72–78
24. Angst J. The course of affective disorders. Psychopathology 1986; 19(suppl 2):47–52
25. Kocsis JH, Croughan JL, Katz MM, et al. Response to treatment with antidepressants of patients with severe or moderate nonpsychotic depression and of patients with psychotic depression. Am J Psychiatry 1990;147: 621–624
26. Croughan JL, Secunda SK, Katz MM, et al. Sociodemographic and prior

clinical course characteristics associated with treatment response in depressed patients. J Psychiatr Res 1988;22:227–237
27. Katz MM, Koslow SH, Maas JW, et al. The timing, specificity and clinical prediction of tricyclic drug effects in depression. Psychosom Med 1987;17:297–309
28. Peselow ED, Sanfilipo MP, Difiglia C, et al. Melancholic/endogenous depression and response to somatic treatment. Am J Psychiatry 1992;149: 1324–1334
29. Kovacs M, Rush AJ, Beck AT, et al. Depressed outpatients treated with cognitive therapy or pharmacotherapy: a one-year follow-up. Arch Gen Psychiatry 1981;38:33–39
30. Blazer D, Hughes DC, George LK. Age and impaired subjective support: predictors of depressive symptoms at one-year follow-up. J Nerv Ment Dis 1992;180:172–178
31. Kaplan HI, Sadock BJ. Comprehensive Textbook of Psychiatry. 6th ed. Baltimore, Md: Williams & Wilkins; 1995
32. Winokur G, Black DW, Nasrallah A. Depressions secondary to other psychiatric disorders and medical illnesses. Am J Psychiatry 1988;145: 233–237
33. Keller MB, Lavori PW, Endicott J, et al. "Double depression": two-year follow-up. Am J Psychiatry 1983;140:689–694
34. Bielski RJ, Friedel RO. Prediction of tricyclic antidepressant response: a critical review. Arch Gen Psychiatry 1976;33:1479–1489
35. Downing RW, Rickels K. Predictors of amitriptyline response in outpatient depressives. J Nerv Ment Dis 1972;154:248–263
36. Hollingshead AB, Redlich FC. Social Class and Mental Illness. New York, NY: Wiley; 1958
37. Duggan C, Lee A, Murray R. Does personality predict long-term outcome in depression? Br J Psychiatry 1990;157:19–24
38. Morgado A, Smith M, Lecrubier Y, et al. Depressed subjects unwittingly overreport poor social adjustment which they reappraise when recovered. J Nerv Ment Dis 1991;179:614–619
39. Greenberg PE, Stiglin LE, Finkelstein SN, et al. The economic burden of depression in 1990. J Clin Psychiatry 1993;54:405–418
40. Klerman GL, Weissman MM. The course, morbidity, and costs of depression. Arch Gen Psychiatry 1992;49:831–834
41. Judd LL. The clinical course of unipolar major depressive disorders. Arch Gen Psychiatry 1997;54:989–991
42. Steketee G, Chambless DL. Methodological issues in prediction of treatment outcome. Clin Psychol Rev 1992;12:387–400
43. Wells K, Burnam M, Rogers W, et al. The course of depression in adult outpatients: results from the Medical Outcomes Study. Arch Gen Psychiatry 1992;49:788–794
44. Hornbrook MC, Goodman MJ. Chronic disease, functional health status, and demographics: a multi-dimensional approach to risk adjustment. Health Serv Res 1996;31:283–307
45. Bindman A, Keane D, Lurie N. Measuring health changes among severely ill patients. Med Care 1990;28:1142–1152
46. Kaplan RM, Anderson JP. The general health policy model: an integrated approach. In: Spilker B, ed. Quality of Life and Pharmacoeconomics in Clinical Trials. 2nd ed. Philadelphia, Pa: Lippincott-Raven; 1996:309–321
47. 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
48. Carroll B, Fielding J, Blashki T. Depression rating scales: a critical review. Arch Gen Psychiatry 1973;28:361–366
49. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–296
50. Stassen HH, Angst J, Delini-Stula A. Severity at baseline and onset of improvement in depression: meta-analysis of imipramine and moclobemide versus placebo. Eur Psychiatry 1994;9:129–136
51. Nierenberg AA, McLean NE, Alpert JE, et al. Early nonresponse to fluoxetine as a predictor of poor 8-week outcome [see comments]. Am J Psychiatry 1995;152:1500–1503
52. Quitkin FM, McGrath PJ, Stewart JW, et al. Chronological milestones to guide drug change: when should clinicians switch antidepressants? Arch Gen Psychiatry 1996;53:785–792
53. First MB, Spitzer RL, Gibbon M, et al. Structured Clinical Interview for DSM-IV Axis I Disorders - Patient Edition (SCID-I/P, Version 2.0). New York, NY: Biometric Research, New York State Psychiatric Institute; 1995
54. Beck AT, Ward CH, Mendelsohn M, et al. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561–571
55. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: 25 years of evaluation. Clin Psychol Rev 1988;8: 77–100
56. Winer B. Statistical Principles in Experimental Design. 2nd ed. New York, NY: McGraw-Hill; 1971
57. Kaplan R, Bush J, Berry C. Health status: types of validity and the index of well-being. Health Serv Res 1976;11:478–507
58. Kaplan RM, Bush JW. Health-related quality of life measurement for evaluation research and policy analysis. Health Psychol 1982;1:61–80
59. Gold M, Siegel J, Russell L, et al, eds. Cost-Effectiveness in Health and Medicine. New York, NY: Oxford University Press; 1996
60. Kaplan R. Hippocratic Predicament: Affordability, Access, and Accountability in Health Care. San Diego, Calif: Academic Press; 1993
61. Kaplan R, Atkins C, Timms R. Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. J Chronic Dis 1984;37:85–95
62. Bombardier C, Ware J, Russell IJ, et al. Auranofin therapy and quality of life in patients with rheumatoid arthritis: results of a multicenter trial. Am J Med 1986;81:565–578
63. Holbrook T, Hoyt D, Anderson J, et al. Functional limitation after major trauma: a more sensitive assessment using the Quality of Well-Being scale: the trauma recovery pilot project. J Trauma 1994;36:74–78
64. Pyne JM, Patterson TL, Kaplan RM, et al. Preliminary longitudinal assessment of quality of life in patients with major depression. Psychopharmacol Bull 1997;33:23–29
65. Hintze JL. NCSS 6.0 Statistical System. User's Guide. Kaysville, Utah: Number Cruncher Statistical Systems; 1998
66. Fleiss JL, Williams JB, Dubro AF. The logistic regression analysis of psychiatric data. J Psychiatr Res 1986;20:195–209
67. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York, NY: John Wiley & Sons; 1989
68. Goodnick PJ, ed. Predictors of Treatment Response in Mood Disorders. Washington, DC: American Psychiatric Press; 1996
69. Lyness JM, Caine ED, Conwell Y, et al. Depressive symptoms, medical illness, and functional status in depressed psychiatric inpatients. Am J Psychiatry 1993;150:910–915
70. Stewart JW, Quitkin FM, McGrath PJ, et al. Social functioning in chronic depression: effects of 6 weeks of antidepressant treatment. Psychiatry Res 1988;25:213–222