The Maudsley Staging Method for Treatment-Resistant Depression: Prediction of Longer-Term Outcome and Persistence of Symptoms

Abebaw Fekadu, M.D.; Sarah C. Wooderson, Ph.D.; Kalypso Markopoulou, M.D.; and Anthony J. Cleare, Ph.D.

Objective: A recently proposed multidimensional method of staging treatment resistance in depression, the Maudsley Staging Method (MSM), has been shown to predict short-term outcome of treatment. This study tested whether the MSM predicts longer-term clinical outcome. We hypothesized that patients with higher scores on the MSM would experience a worse longer-term outcome in terms of time spent in a depressive episode and level of functional impairment.

Method: From May through July of 2008, we followed up patients with treatment-resistant depression discharged from an inpatient unit of an affective disorders service; all had MSM scores previously calculated from preadmission clinical data. We used the Longitudinal Interval Follow-up Evaluation (LIFE) chart to determine the monthly symptomatic course of depression blind to initial MSM scores. We employed a regression model to adjust for various confounding factors, including variable duration of follow-up, to determine the independent association of MSM scores with persistence of depressive disorder.

Results: We assessed 62 of 80 eligible patients (78%) in a median follow-up duration (interquartile range) of 29.5 (19.0–52.5) months. The MSM independently predicted (1) being in an episode for 50% or longer of the follow-up duration (OR = 2.11, 95% CI = 1.25 to 3.57), (2) being in an episode at the time of follow-up assessment (OR = 1.89, 95% CI = 1.17 to 3.05), (3) being persistently in an episode throughout the follow-up period (OR = 2.01, 95% CI = 1.14 to 3.54), and (4) total months spent in a depressive episode (OR = 1.22, 95% CI = 1.06 to 1.40). The MSM also predicted functional impairment. Antidepressant count and the Thase and Rush model did not independently predict persistence of depression or functional impairment.

Conclusion: The MSM appears to have reasonable predictive validity regarding the longer-term course of illness, particularly persistence of depressive episodes. The MSM may be a useful, and possibly an improved, alternative to existing models of staging of treatment-resistant depression.

J Clin Psychiatry 2009;70(7):952–957

Received Sept. 20, 2008; accepted Jan. 2, 2009. From the Institute of Psychiatry, Section of Neurobiology of Mood Disorders, Department of Psychological Medicine and Psychiatry, King's College London, United Kingdom.

Drs. Fekadu, Wooderson, Markopoulou, and Cleare report no financial affiliations or other relationships relevant to the subject of this article.

Corresponding author and reprints: Anthony J. Cleare, Ph.D., King's College London, Institute of Psychiatry, Section of Neurobiology of Mood Disorders, Department of Psychological Medicine and Psychiatry, 103 Denmark Hill, London SE5 8AZ, United Kingdom (e-mail: A.Cleare@iop.kcl.ac.uk).

The personal and societal impact of depression in terms of causing personal suffering, disability, and financial costs is undoubtedly substantial.[1–4] Because of its relatively common occurrence and inevitably longer duration and greater severity,[5–8] treatment-resistant depression (TRD) is likely to make a major contribution to the overall impact produced by depression.[5,9,10]

In order to alleviate the burden caused by depression, TRD has to be a priority area for pragmatic research. However, the lack of a uniformly accepted and valid definition of TRD and the absence of well-validated methods to stage the severity of TRD present significant obstacles to progress in this area.

The commonly used criterion to stage treatment resistance has been the Thase and Rush model.[11] This model is relatively simple to use, but its hierarchical nature[12] and the restricted rating options limit its usefulness. Given the large number of treatment options currently available and the limited flexibility of the model, the Thase and Rush model should probably be considered outdated.

The Massachusetts General Hospital staging method (MGH-S)[13] has become a useful alternative. This quantitative method of staging primarily relies on the number of antidepressant medications used to estimate level of resistance. Intensity of treatment and augmentation strategies are also considered, and a special weight is given for failure of treatment with electroconvulsive therapy (ECT). The model allows flexibility to incorporate as many medications as required in gauging degree of treatment resistance. However, with the potential for a large number of treatment options currently available, the system may be less efficient and less discriminating. Thus, data obtained may not inform intervention strategies or fully enhance understanding and

communication. There is also no clear evidence supporting the magnitude of the special weight given to treatment with ECT. There is limited evidence on the utility of this model.[14]

The European method of staging relies on matching treatment resistance to the specific class of medication used.[15] This method is useful in as far as it recognizes the role of duration of illness in treatment resistance. The model is limited in its scope and makes unwarranted assumptions regarding the differential effectiveness of antidepressant medications. One further staging model[10] was recently explored and has been shown to have some validity on cross-sectional assessment. This model was based on depressive subtypes on a dimension of severity (psychotic, melancholic, and nonmelancholic). Again, this is of limited scope.

Given the complexity of psychiatric disorders, staging resistant depression on the basis of a single factor (treatment response) is not satisfactory.[11,15] We recently proposed a multidimensional staging method, the Maudsley Staging Method (MSM), based on available literature and empirical data from a sample of patients with TRD.[16] In addition to number of failed treatment trials, this model incorporates additional factors considered to be closely related to the depressive illness itself: severity and duration of the presenting episode (Appendix 1).

The MSM allows the actual stage of treatment resistance to be represented as a single score, which may vary between 3 and 15. It also allows summarizing the stage of resistance into 3 ordinal categories (mild, moderate, and severe). The validity of the method in predicting outcome at discharge from hospital was also tested on the basis of empirical data. For deriving the initial empirical data, we extracted relevant information from the medical notes of 88 patients discharged from an inpatient unit of an affective disorders service. These cases were selected on the basis of their having complete information that would allow accurate scoring of the level of treatment resistance according to the MSM. Analysis of the data showed that patients with a higher MSM score were more likely to be symptomatic when discharged—rather than in remission.[16] For the present study, we contacted these same patients again to determine whether MSM would predict longer-term persistence of depression. We hypothesized that MSM would predict persistence of depressive disorder, and specifically we hypothesized that those with a higher score on the MSM would be more likely to experience a persistent depressive episode, to spend most of the follow-up time in a depressive episode, to be in a depressive episode, and to exhibit functional impairment at the time of follow-up assessment.

METHOD

Study Sample

The study was conducted in a tertiary research center in the United Kingdom. The base population consisted of patients with confirmed TRD, defined by failure of at least 1 antidepressant medication, and discharged from the inpatient unit at least 1 year prior to follow-up. We recruited patients for whom we had initially computed MSM scores to provide the initial empirical evidence to support our proposed staging model.[16] Follow-up assessments were conducted between May 2008 and July 2008. The study was approved by the Joint South London and Maudsley and the Institute of Psychiatry National Health Service Research Ethics Committee, and all participants provided informed consent.

Assessment

The key assessment instrument was the Longitudinal Interval Follow-up Evaluation (LIFE) chart.[17] The LIFE chart is a well-established follow-up evaluation scale that allows the weekly or monthly symptomatic state of a patient to be rated retrospectively at intervals of 6 months or longer. Symptoms are ordinarily rated on a 6-point scale, called the Psychiatric Status Rating (PSR), which ranges from an asymptomatic state (score of 1) to a severe episode with psychosis or severe impairment (score of 6). Although operationally linked to the Research Diagnostic Criteria[18] in the original design, the PSR has been subsequently adapted for use with the *Diagnostic and Statistical Manual of Mental Disorders*.[19] In this study, we used the modified PSR rating as adapted for a follow-up study of patients with depression.[20,21] The modification expands the PSR ratings from 6 to 7: a score of 1 or 2 corresponds to remission, a score of 3 or 4 corresponds to residual symptoms, and a score of 5, 6, or 7 corresponds to being in a depressive episode (mild = 5, moderate = 6, severe = 7). The LIFE chart was completed by 2 research psychiatrists, and the PSRs recorded the monthly clinical status of patients. Similar monthly ratings have been made in other studies.[20,21]

The cross-sectional symptomatic state at the time of follow-up was ascertained using the Quick Inventory of Depressive Symptomatology, Clinician-Rated (QIDS-C) and Self-Report (QIDS-SR) versions.[22] Both instruments contain identical items, with DSM-IV criterion symptoms forming the core of the instruments. These instruments have been established to have good psychometric properties.[22]

We also administered the Global Assessment of Functioning (GAF) scale,[23] a widely used dimensional scale that allows rating of impairment in psychological, social, and occupational functioning. The GAF scores vary from 0 (rated when there is no adequate information to complete the scale) to 100 (rated for exceptionally "superior functioning"). Although the GAF rates for symptom severity and functional impairment, our focus was on functional impairment.

Data Management

Data were initially entered using the Statistical Package for the Social Sciences (SPSS) for Windows, Release 15

(SPSS Inc., Chicago, Ill.). Subsequent data analyses were carried out using SPSS and Stata for Windows, version 10 (StataCorp, College Station, Tex.).

Outcome variables considered were proxy indices for persistence of depressive disorder. When applicable, both continuous and dichotomous (binary) outcomes were considered. Dichotomous outcomes included having a persistent episode throughout the period of follow-up (versus not having a persistent episode), being in an episode for 50% or longer of the follow-up duration, and being in an episode and experiencing functional impairment at the time of follow-up assessment. All PSR scores of 5 through 7 were equated to a depressive episode, and the time frame for this was 1 month as prescribed in the LIFE rating for this study. Functional impairment was primarily defined as having a GAF score below 65, which is indicative of at least mild, but significant, functional impairment. The main continuous outcome was the number of months spent in a depressive episode during follow-up. We also looked at symptomatic state (QIDS scores) at time of follow-up.

The main analytic method employed was binary logistic regression using SPSS. We also employed exact logistic regression using Stata when the sample size in a cell was small. Because of the lack of normal distribution and the inflated number of zero scores, we employed the zero-inflated negative binomial regression method as implemented in the Stata statistical program to analyze data on months spent in an episode. Analysis was carried out in 3 stages. The first stage was fitting a univariate (unadjusted) model. In the second stage, the MSM or other applicable variables—and only factors found to be associated with outcome in the univariate model, or because of relevance for the analysis—were entered into the model (for example, while computing the association of MSM with months spent in an episode, duration of follow-up was a key element and was entered into the model in the second stage). In the third stage, all factors considered to be relevant in predicting outcome were fitted. These adjustments are shown in Tables 2 through 4.

To assess for any significant difference between patients that were and were not followed up despite being in the initial cohort, we compared the 2 groups for sociodemographic and clinical characteristics, including level of depressive symptoms at discharge, measured using the Hamilton Rating Scale for Depression (HAM-D).[24]

Table 1. Sociodemographic and Clinical Characteristics of the Cohort (62 patients with treatment-resistant depression)

Characteristic	Value
Baseline characteristics	
Sex, %	
Female	71.0
Marital status, %	
Single	24.2
Married	61.3
Postmarital[a]	14.5
MSM model summary, %	
Mild	3.2
Moderate	32.3
Severe	64.5
Thase and Rush score summary, %	
< 5	16.1
5	83.9
Age, mean (SD), y	49.6 (13.3)
Years of education, mean (SD)	13 (3.4)
Duration of current episode, median (IQR), y	4.0 (1.7–7.5)
MSM score, mean (SD)	10.8 (2.1)
Follow-up characteristics	
In episode at follow-up (PSR of 5–7), %	
Yes	32.3
No	67.7
Persistent depression (PSR of 5–7 throughout), %	21.0
In episode (PSR of 5–7) for 50% or longer of the follow-up duration, %	37.7
Functional impairment (GAF score < 65) at follow-up, %	41.9
QIDS-C score, mean (SD)	7.2 (5.3)
Percent of time in episode, median (IQR)	27.5 (0.0–80.3)
Months in follow-up, median (IQR)	29.5 (19.0–52.5)

[a]Includes separated, divorced, and widowed.
Abbreviations: GAF = Global Assessment of Functioning; IQR = interquartile range; MSM = Maudsley Staging Method; PSR = Psychiatric Status Rating; QIDS-C = Quick Inventory of Depressive Symptomatology, Clinician-Rated.

RESULTS

Sample Characteristics and Summary of Clinical Outcomes

From the original 88 cases forming the basis of the initial empirical data for the MSM, 6 individuals were deceased and 2 had a follow-up duration shorter than 1 year and were excluded. Of the 80 potentially eligible people who were approached for this study, 62 with TRD were successfully interviewed (follow-up rate, 78%). There were no significant differences in sociodemographic characteristics (sex, age, and years of education) or relevant clinical characteristics (duration of presenting depressive episode, severity of illness, score on the HAM-D at discharge, and severity of illness) between patients who were and were not interviewed. However, there was a small but significantly higher MSM score among those who were interviewed (mean difference in score = 1.01; p value = .03). A summary of sample characteristics is shown in Table 1. Following discharge, patients were followed up for a median duration (interquartile range) of 29.5 (19.0–52.5) months, with a mean (SD) of 36.4 (21.8) months.

Most of the patients who took part in the follow-up study were women (71.0%), with a mean (SD) age of 49.6 (13.3) years. Most had severe treatment resistance—64.5% according to the MSM and 83.9% according to the Thase and Rush model. The median duration (interquartile range) of the index episode of depression at baseline was 4.0 (1.7–7.5) years. More than one third of patients (37.7%) had been in a depressive episode for 50% or longer of the follow-up period, while one fifth (21.0%) were continuously in an episode. At the time of follow-up assessment, about one third of patients (32.3%) were in a depressive episode, while 41.9% were functionally impaired.

Table 2. Prediction of Persistence of Depressive Disorder Using the Maudsley Staging Method (MSM) Adjusted for Relevant Clinical and Sociodemographic Factors

Outcome/Model	Odds Ratio	95% CI	p Value
In an episode for at least 50% of the follow-up period			
MSM score (crude)	1.48	1.05 to 2.09	.024
Model 1[a]	1.64	1.12 to 2.40	.010
Model 2[b]	2.11	1.25 to 3.57	.005
Persistent episode throughout the follow-up period			
MSM score (crude)	1.85	1.23 to 2.78	.003
Model 1[c]	1.74	1.13 to 2.68	.012
Model 2[d]	2.01	1.14 to 3.54	.015
Months spent in an episode[e]			
	Coefficient		
MSM score (crude)	1.08	0.89 to 1.32	.413
Model 1[f]	1.18	1.03 to 1.36	.017
Model 2[g]	1.17	1.04 to 1.32	.008
Model 3[h]	1.22	1.06 to 1.40	.006

[a]Adjusted for age and years of education.
[b]Additionally adjusted for sex, marital status, duration of hospital admission, and duration of follow-up.
[c]Adjusted for years of education.
[d]Additionally adjusted for age, sex, marital status, duration of hospital admission, and duration of follow-up.
[e]All models used the binary logistic regression method except this outcome, which used a zero-inflated negative binomial regression method.
[f]Adjusted for duration of follow-up.
[g]Also adjusted for age and duration of admission.
[h]Final model additionally adjusted for sex, marital status, and years of education.

Table 3. Prediction of Clinical and Functional Status at the End of Follow-Up Using the Maudsley Staging Method (MSM) Adjusted for Relevant Clinical and Sociodemographic Factors

Outcome/Model[a]	Odds Ratio	95% CI	p Value
In an episode at time of final interview			
MSM score (crude)	1.46	1.06 to 2.02	.021
Model 1[b]	1.71	1.11 to 2.63	.015
Model 2[c]	1.89	1.17 to 3.05	.010
Functional impairment at follow-up			
MSM score (crude)	1.57	1.14 to 2.15	.005
Model 1[d]	1.48	1.06 to 2.06	.021
Model 2[e]	1.55	1.08 to 2.22	.017

[a]All models used the binary logistic regression method.
[b]Adjusted for age and years of education.
[c]Additionally adjusted for sex, marital status, duration of hospital admission, and duration of follow-up.
[d]Adjusted for years of education.
[e]Additionally adjusted for age, sex, marital status, duration of hospital admission, and duration of follow-up.

Prediction of Outcome

In the fully adjusted model, higher odds of prediction for higher MSM score were found for the occurrence of persistent depressive episode throughout follow-up (OR = 2.01, 95% CI = 1.14 to 3.54, p = .015) and for the occurrence of depressive episodes lasting for 50% or longer of the follow-up period (OR = 2.11, 95% CI = 1.25 to 3.57, p = .005) (Table 2). The odds of finding a depressive episode at follow-up assessment were also nearly 2-fold (OR = 1.89, 95% CI = 1.17 to 3.05, p = .010) (Table 3). A unit increase in the MSM score predicted spending about one extra week in a depressive episode (model coefficient = 1.22, 95% CI = 1.06 to 1.40, p = .006) (Table 2). Functional impairment, defined as having a GAF score of less than 65, was also predicted by the MSM (OR = 1.55, 95% CI = 1.08 to 2.22, p = .017) (Table 3).

The MSM also predicted symptom severity as measured with the QIDS-C and QIDS-SR at the time of follow-up assessment. Thus, for prediction of the QIDS-SR, the adjusted standardized coefficient was 0.50 (95% CI = 0.25 to 3.24, p = .024), and, for the QIDS-C, it was 0.29 (95% CI = 0.02 to 20.37, p = .046). Summary scores of the MSM also predicted persistence of depressive disorder and functional impairment. For being functionally impaired, the adjusted odds ratio was 4.75 (95% CI = 1.12 to 20.37, p = .036), while, for being in episode for 50% or longer of the follow-up period, it was 12.21 (95% CI = 1.61 to 92.63).

Other staging methods and individual staging components were assessed for prediction of the same clinical and functional outcome domains (Table 4). None were predictive in the fully adjusted model other than longer duration of the index depressive episode, which predicted being in a depressive episode for 50% or longer of the follow-up duration (OR = 1.22, 95% CI = 1.02 to 1.46, p = .031). The Thase and Rush model also failed to predict outcome. However, most patients in the sample had the maximum score of 5, and this was particularly the case among those with a persistent episode. Number of antidepressant medications used and varying weights for ECT did not predict outcome.

DISCUSSION

As part of our initial proposal of a new multidimensional staging model of treatment resistance—the Maudsley Staging Method—we had shown the utility of the proposed model in predicting nonremission following specialist inpatient treatment.[16] To our knowledge, the only other staging method that showed validity in predicting nonremission following treatment was the Massachusetts General Hospital staging method.[14] In the same study, that group found that the Thase and Rush model failed to predict this outcome.[14] Another report attempted to compare the concurrent validity of a symptom-based staging method[10] to clinicians' ratings of treatment resistance. Although the method was not a formally recommended method of staging, the symptom-based grouping of depression was predictive of clinicians' rating of treatment resistance.[10] Our report here takes the validation of staging methods a step further by looking at whether the methods can predict the persistence of depressive illness in both the shorter- and the longer-term course of illness.

In this follow-up study spanning 1 to 7 years following discharge from a specialist inpatient unit, we show that, as hypothesized, higher MSM scores were associated with persistence of depressive disorder. This association was shown for longitudinal course indicators of persistence of disorder

  © Copyright 2009 Physicians Postgraduate Press, Inc.

Table 4. Prediction of Persistence of Depressive Illness and Functional Impairment (as indices of resistance) Using Alternative Models or Model Components of Staging Treatment Resistance in Treatment-Resistant Depression

	Binary Outcomes and Model Estimates[a]							
	In Current Episode		In Persistent Episode		In Episode for 50% or Longer of Follow-Up Duration		Functional Impairment	
Model Variable	OR (95% CI)	p Value	OR (95% CI)	p Value	OR (95% CI)	p Value	OR (95% CI)	p Value
Thase and Rush stages	3.38 (0.79 to 14.53)	.102	2.15 (0.85 to NUB)	.122	2.16 (0.75 to 6.27)	.156	1.49 (0.74 to 3.03)	.269
Antidepressant count	2.29 (0.88 to 5.98)	.090	1.48 (0.69 to 3.17)	.32	1.58 (0.81 to 3.10)	.180	1.90 (0.95 to 3.82)	.071[b]
Antidepressant count, mood stabilizer, and ECT	1.52 (0.97 to 2.38)	.065	1.39 (0.84 to 2.30)	.202	1.51 (0.97 to 2.36)	.068	1.73 (0.97 to 3.08)	.061[c]
Antidepressant count, mood stabilizer, and ECT (ECT given weighted score of 2 points)	1.39 (0.95 to 2.05)	.090	1.36 (0.87 to 2.14)	.182	1.39 (0.95 to 2.03)	.087	1.49 (0.92 to 2.40)	.107[d]
Antidepressant count, mood stabilizer, and ECT (ECT given weighted score of 3 points)	1.29 (0.93 to 1.80)	.124	1.32 (0.88 to 1.97)	.180	1.30 (0.94 to 1.79)	.111	1.41 (1.00 to 1.99)	.054
Duration of current episode	1.01 (0.97 to 1.22)	.152	1.08 (0.96 to 1.21)	.20	1.22 (1.02 to 1.46)	.031[e]	1.10 (0.97 to 1.24)	.120

[a]All values are unadjusted except where specified; where specified, adjustments were made for age, sex, years of education, duration of follow-up, and duration of admission in inpatient unit at index admission.
[b]Adjusted (unadjusted model: OR = 2.09, 95% CI = 1.01 to 4.34, p = .047).
[c]Adjusted (unadjusted model: OR = 1.73, 95% CI = 1.08 to 2.78, p = .024).
[d]Adjusted (unadjusted model: OR = 1.55, 95% CI = 1.03 to 2.33, p = .036).
[e]Adjusted (unadjusted model: OR = 1.20, 95% CI = 1.03 to 1.41, p = .034).
Abbreviations: ECT = electroconvulsive therapy, NUB = no upper bound.

(i.e., more months spent in episode, experiencing a single persistent episode, or being in an episode for 50% or longer of the follow-up duration) as well as for symptomatic state at time of assessment. This association also extended to functional impairment at the time of the follow-up interview.

These findings, taken together with the findings in the previous report of the prediction of discharge outcome,[16] suggest that the MSM not only is theoretically meaningful but may also have utility in clinical practice and research. The only other model we formally tested in this study for prediction was the Thase and Rush model, which was not found to be predictive of persistence of disorder. However, this negative finding may be partly because most cases in the study had severe resistance, with a maximum score of 5 according to the Thase and Rush model. Thus, this particular model may be lacking the ability to discriminate different levels of treatment resistance in this specialist patient population and, instead, may be better suited to measure resistance in community settings or nonspecialist services, although this suggestion remains to be tested empirically. As discussed above, the main limitations of the Thase and Rush method are its lack of flexibility and its restricted rating options, and it may have to be considered outdated. Although we did not formally test other models for prediction of outcome, treatment count did not predict persistence of disorder. This lack of predictive validity was not significantly affected by scoring augmentation strategies, treatment with ECT, and antidepressant treatment count together. Varying the weight for ECT also did not change the prediction. This finding suggests that a unidimensional staging model of resistance that uses treatment trials as the sole criterion may be a less useful model.

Our sample is unusual in that patients were from a tertiary unit, with severe treatment resistance and, thus, not necessarily representative of the larger population of patients with TRD. Nevertheless, the original staging method was developed using the same sample, and the present findings suggest consistency in that the factors contained within the MSM are consistent in predicting both shorter- and longer-term outcome. The sample size was also relatively small, but this was because of the necessity for using a constrained sample from the outset, as we relied on a previous sample with information on MSM. Again, the findings are strong and consistent for both cross-sectional clinical and functional outcomes at follow-up as well as for the longitudinal course-of-illness data. This consistency indicates the strength of the findings despite the small sample size. Some of the data were collected retrospectively, specifically the LIFE chart data, raising the possibility of recall bias. However, all available information was used for these ratings; thus, in addition to patient interviews, clinical records and information from carers were used. LIFE chart data collection is a validated methodology,[17,25] and we avoided observer bias by conducting all assessments blind to the initial MSM score.

Overall, despite some of the limitations mentioned above, this study indicates that the Maudsley Staging Method may be a useful model for staging the degree of treatment resistance in depression with relatively clear clinical and research utility. Nevertheless, our findings should be considered preliminary given the unusual study setting and small sample size. We recommend further assessment of the utility of this staging model based on a prospective study design, a larger sample size, and broader study settings.

REFERENCES

1. Greenberg PE, Kessler RC, Birnbaum HG, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry.* 2003;64(12):1465–1475.
2. Murray CJL, Lopez AD. *The Global Burden of Disease: a Comprehensive Assessment of Mortality and Disability From Disease, Injuries, and Risk Factors in 1990 and Projected to 2020*. Cambridge, Mass: Harvard University Press; 1996.
3. Thomas CM, Morris M. Cost of depression among adults in England in 2000. *Br J Psychiatry.* 2003;183:514–519.
4. World Health Organization. *The World Health Report 2001–Mental Health: New Understanding, New Hope*. Geneva, Switzerland: World Health Organization Press; 2001.
5. Greden JF. The burden of disease for treatment-resistant depression. *J Clin Psychiatry.* 2001;62(suppl 16):26–31.
6. Souery D, Papakostas GI, Trivedi MH. Treatment-resistant depression. *J Clin Psychiatry.* 2006;67(suppl 6):16–22.
7. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry.* 2006;163(1):28–40.
8. Ustun TB, Kessler RC. Global burden of depressive disorders: the issue of duration. *Br J Psychiatry.* 2002;181:181–183.
9. George MS, Rush AJ, Marangell LB, et al. A one-year comparison of vagus nerve stimulation with treatment as ususal for treatment-resistant depression. *Biol Psychiatry.* 2005;58(5):364–373.
10. Malhi GS, Parker GB, Crawford JG, et al. Treatment-resistant depression: resistant to definition? *Acta Psychiatr Scand.* 2005;112(4):302–309.
11. Thase ME, Rush AJ. When at first you don't succeed: sequential strategies for antidepressant nonresponders. *J Clin Psychiatry.* 1997;58 (suppl 13):23–29.
12. Parker GB, Malhi GS, Crawford JG, et al. Identifying "paradigm failures" contributing to treatment-resistant depression. *J Affect Disord.* 2005;87(2–3):185–191.
13. Fava M. Diagnosis and definition of treatment-resistant depression. *Biol Psychiatry.* 2003;53(8):649–659.
14. Petersen T, Papakostas GI, Posternak MA, et al. Empirical testing of two models of staging antidepressant treatment resistance. *J Clin Psychopharmacol.* 2005;25(4):336–341.
15. Souery D, Amsterdam J, de Montigny C, et al. Treatment resistant depression: methodological overview and operational criteria. *Eur Neuropsychopharmacol.* 1999;9(1–2):83–91.
16. Fekadu A, Wooderson S, Donaldson C, et al. A multidimensional tool to quantify treatment resistance in depression: the Maudsley Staging Method. *J Clin Psychiatry.* 2009;70(2):177–184.
17. Keller MB, Lavori W, Friedman B, et al. The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry.* 1987;44(6):540–548.
18. Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria: rationale and reliability. *Arch Gen Psychiatry.* 1978;35(6):773–782.
19. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition. Washington, DC: American Psychiatric Association; 1994.
20. Kennedy N, Abbott R, Paykel ES. Remission and recurrence of depression in the maintenance era: long-term outcome in a Cambridge cohort. *Psychol Med.* 2003;33(5):827–838.
21. Kennedy N, Paykel ES. Residual symptoms at remission from depression: impact on long-term outcome. *J Affect Disord.* 2004;80(2–3):135–144.
22. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry.* 2003;54(5):573–583.
23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition. Washington, DC: American Psychiatric Association; 1994:32.
24. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23(1):56–62.
25. Mueller TI, Leon AC, Keller MB, et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *Am J Psychiatry.* 1999;156(7):1000–1006.

Appendix 1. Maudsley Staging Parameters and Suggested Scoring Conventions[a]

Parameter/Dimension	Parameter Specification	Score
Duration	Acute (≤ 12 months)	1
	Sub-acute (13–24 months)	2
	Chronic (> 24 months)	3
Symptom severity (at baseline)	Subsyndromal	1
	Syndromal	
	Mild	2
	Moderate	3
	Severe without psychosis	4
	Severe with psychosis	5
Treatment failures		
Antidepressants	Level 1: 1–2 medications	1
	Level 2: 3–4 medications	2
	Level 3: 5–6 medications	3
	Level 4: 7–10 medications	4
	Level 5: > 10 medications	5
Augmentation	Not used	0
	Used	1
Electroconvulsive therapy	Not used	0
	Used	1
Total		(15)

[a]Reprinted with permission from Fekadu et al.[16]

© Copyright 2009 Physicians Postgraduate Press, Inc.