Validity of the Maudsley Staging Method in Predicting Treatment-Resistant Depression Outcome Using the Netherlands Study of Depression and Anxiety

Sjoerd M. van Belkum, MD[a,*]; Hanneke H. Geugies, MSc[a]; Thom S. Lysen, MD[b]; Anthony J. Cleare, PhD[c]; Frenk P. M. L. Peeters, MD, PhD[d]; Brenda W. J. H. Penninx, PhD[e]; Robert A. Schoevers, MD, PhD[a]; and Eric G. Ruhe, MD, PhD[a]

ABSTRACT

Objective: We investigated if the degree of treatment resistance of depression, as measured by the Maudsley Staging Method (MSM), is predictive of a worse depression outcome by using a large naturalistic cohort of depressed patients.

Methods: 643 subjects from the general population, primary care, and secondary care who suffered from current depressive disorder were included from the Netherlands Study of Depression and Anxiety baseline assessment. The diagnostic criterion was major depressive disorder (MDD) in the last month, based on the Composite Interview Diagnostic Instrument (CIDI), or a CIDI diagnosis of MDD in the past 6 months with an Inventory of Depressive Symptomatology Self-Report score > 24 at baseline. In these subjects, composite scores of the MSM, based on duration, severity, and treatment history of current episode, were determined retrospectively. We then determined if the MSM score prospectively predicted the 2-year course of depression after baseline. The primary outcomes were percentage of follow-up time spent in a depressive episode and being "mostly depressed" (≥ 50% of the follow-up) between baseline and 2-year follow-up.

Results: The MSM predicted "percentage of follow-up time with depression" ($P < .001$) and was associated with being "mostly depressed" (OR = 1.40; 95% CI, 1.23–1.60; $P < .001$). These effects were not modified by having received treatment.

Conclusions: The current study shows that the MSM is a promising tool to predict worse depression outcomes in depressed patients. In this study that adds to previous work, we show the applicability of MSM in a wider range of primary and secondary care patients with depression.

J Clin Psychiatry 2018;79(1):17m11475

To cite: van Belkum SM, Geugies HH, Lysen TS, et al. Validity of the Maudsley Staging Method in predicting treatment-resistant depression outcome using the Netherlands Study of Depression and Anxiety. *J Clin Psychiatry.* 2018;79(1):17m11475.

To share: https://doi.org/10.4088/JCP.17m11475

[a]Department of Psychiatry, Research School of Behavioural and Cognitive Neurosciences (BCN), Interdisciplinary Center Psychopathology of Emotion Regulation (ICPE), University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[b]Department of Epidemiology, University Medical Center Rotterdam, Erasmus University, Rotterdam, The Netherlands

[c]Department of Psychological Medicine and Psychiatry, Centre for Affective Disorders, Institute of Psychiatry, King's College London, London, United Kingdom

[d]Department of Psychiatry and Neuropsychology, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands

[e]Department of Psychiatry, Amsterdam Public Health and Amsterdam Neuroscience, VU University Medical Center, Amsterdam, The Netherlands

**Corresponding author:* Sjoerd M. van Belkum, MD, Universitair Centrum Psychiatrie CC 30, University Medical Center Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands (s.m.van.belkum@umcg.nl).

Treatment of major depressive disorder (MDD) mainly consists of different forms and combinations of psychotherapy and antidepressant medication. Overall, such treatment has moderate efficacy.[1–4] However, treatment appears to be ineffective for a particular group of patients, who are then categorized as suffering from treatment-resistant depression (TRD). In the largest treatment study to date, the Sequenced Treatment Alternatives to Relieve Depression (STAR*D), 49% of patients showed a response (≥ 50% improvement on the Quick Inventory of Depressive Symptomatology–Self-Report [QIDS-SR_{16}]) and 37%, remission (≤ 5 on the QIDS-SR_{16}) after the first antidepressant.[5] Remission rates gradually declined with each sequential step thereafter. Moreover, in that study, even after 4 treatment trials, 33% of patients had not achieved remission.[5] Treatment resistance is the main cause for the large societal costs of depression.[6,7] Timely identification of patients with treatment resistance would provide the opportunity of an earlier start of intensified treatment regimens to address MDD symptoms more aggressively and secure potentially better health care outcomes.

Unfortunately, research on TRD is hampered by the lack of consensus on its definition. Treatment-resistant depression is often categorically defined as nonresponse to ≥ 2 adequate antidepressants trials.[8–12] However, over 10 other definitions of TRD have been proposed, differing mostly on the number of pharmacologic treatment steps patients have had.[9,10,13] Furthermore, although TRD is mostly represented as a dichotomy, this characterization does not seem to represent clinical reality, as was shown in the STAR*D and other antidepressant-switch trials.[5,14] Treatment-resistant depression might therefore be better considered as a dimensional construct.[8,9] Treatment resistance, then, is scored on a spectrum, running from quick remission (sometimes even without treatment) to the other extreme: severe treatment resistance when no treatment response occurs after electroconvulsive therapy and other third-line treatment regimens.

Over the last decade, progress has been made in methods to quantify TRD and use this quantification to predict the course and outcome of depression.[8] However, these methods have been validated to a limited extent only. Of these methods, the Maudsley Staging Method (MSM) appeared to be one of the most promising.[8,15] The

Clinical Points

- The Maudsley Staging Method has been validated as a tool to quantify and predict depression course and outcome, but generalizability to a much larger community-based population is required to maximize its utility.
- This study shows that the Maudsley Staging Method can be a reliable and valid tool to predict poor outcome in depressed persons with respect to both a natural course without treatment and a subsequent course of treatment according to contemporary guidelines.

MSM was created to represent the broad theoretical basis of treatment resistance and is aimed at predicting outcome of depression. During the development of the MSM, incorporation of severity and duration in predicting worse depression outcome showed added value, as these are strong and consistent predictors of the prognosis of MDD.[16–18] Both the MSM as a whole and its different components were shown to independently predict both failure to achieve remission[15] and persistence of the depressive episode.[19]

However, the MSM has been investigated using only a relatively small sample (N = 88) of patients who were treated in tertiary care.[15,19] Generalizability to the much larger community-based population of depressed patients and those attending primary and secondary care is required to maximize the utility of the tool for predicting remission, episode persistence, and/or future treatment resistance. Therefore, the aim of this study was to further validate the predictive value of the MSM. We examined if the degree of treatment resistance over its full spectrum, as measured by the MSM, is predictive of a chronic course of illness by using the large naturalistic cohort of the Netherlands Study of Depression and Anxiety (NESDA).[20] We expected the MSM to be predictive of the longitudinal course of illness during 2 years of follow-up.

METHODS

Setting

The NESDA is a multisite, naturalistic cohort study with data from 2,329 patients with MDD and/or anxiety, sampled from the general population (by interviewing members of private households or children of parents who were treated for depressive disorder), primary care (ie, general practitioner), and secondary care (ie, specialized mental health institutions), and 652 controls, aged 18 through 65 years.[20] After approval from the Medical Ethics Review Committee of the VU University Medical Center Amsterdam, written informed consent of every subject was obtained.

Sample

Inclusion criteria for our study were (1) a diagnosis of MDD in the last month (based on the Composite Interview Diagnostic Instrument [CIDI], lifetime version 2.1)[21] *or* a CIDI diagnosis of MDD in the past 6 months with an Inventory of Depressive Symptomatology Self-Report (IDS-SR) score > 24 (the clinical cutoff value for moderately severe depression[22,23]) at baseline; (2) availability of all data needed to calculate the MSM score; and (3) availability of sufficient data to determine outcome during 2 years of follow-up. To cover the full spectrum of treatment resistance, from null to a more severe form, we also included depressed subjects from primary or secondary care who had not yet received treatment, as well as subjects from the general population who, despite having depressive symptoms, had not yet sought treatment.

Determinants: MSM

The MSM is composed of 3 items: (1) duration, which is scored 1 to 3; (2) severity, which is scored 1 to 5; and (3) treatment failures. Treatment failures are scored 0 to 5 with regard to antidepressants used in the current episode, 0 or 1 with regard to augmentation used in the current episode, and 0 or 1 with regard to electroconvulsive therapy used in the current episode.[15] (See Supplementary Methods, eTable 1, for a reprint of the MSM published by Fekadu et al.[15])

We used different variables from the NESDA database to obtain the 3 item scores to determine the degree of treatment resistance.

1. Duration of the current episode at baseline was established using the retrospective Life-Chart Interview (LCI).[24] The LCI relies on self-generated and affectively laden landmarks as anchors for participants to refresh memory. After these anchors were determined, presence and severity of depressive symptoms were assessed during each quarter of the past 4 years prior to baseline.
2. Severity of depression was assessed according to *DSM-IV*, as determined by the CIDI.
3. Treatment history was scored based on the amount of subsequently used antidepressants and augmentation strategies during the index episode, at, and prior to baseline. A specific drug was scored as being used if the frequency of use was daily, if the dosage was at least the daily defined dose, and if it was used for at least 4 weeks (1 month).[25] (See also Supplementary Methods.)

The subscores of these 3 items (duration, severity, and total score of treatment failures) are added together to obtain a total score.

Outcome: Course Trajectory of Depression in NESDA

In the present article, following Fekadu et al,[19] we focused on the intensity and duration of depressive symptoms during 2-year follow-up in subjects with a depressive disorder (index episode) at baseline. In order to predict the course of the depressive episode after baseline, the primary outcome was persistence of the depressive episode based on LCI data between baseline and 2-year follow-up. We made 2 different variables: (1) The variable "percentage of follow-up time

with depression" was expressed as the ratio between months spent in a depressive episode since baseline until remission, divided by total follow-up time (24 months). In line with the prevailing method in the NESDA database,[26] remission was defined as experiencing a period of 3 consecutive months without symptoms, or with symptoms but without burden or interference with life (as indicated by the participant). The month of remission was defined as the first month after this 3-month period. (2) Analogous to the previous validation study,[19] we defined the categorical variable persistent depression as being persistently depressed for ≥ 50% of the time of our follow-up period of 2 years.

For our secondary outcome, we used course trajectories as described in NESDA by Rhebergen et al.[27] Rhebergen and colleagues used latent class growth analysis, a statistical data-driven technique to describe patterns inherently present in data, in this case representing depression course trajectories. In brief, with input of LCI data from NESDA Wave 3, which covers the entire 2-year follow-up period, 5 course trajectories were identified: (1) a quick remission course, (2) a decline course with moderate severity, (3) a decline course with high severity, (4) a chronic course with moderate severity, and (5) a chronic course with high severity.[27]

Statistical Analysis

Analyses were performed with IBM-SPSS, version 20 (IBM, Chicago, Illinois). Analyses for primary outcomes were performed using linear regression analysis and logistic regression analysis for "percentage time depressed" and "persistent depression," respectively. For our secondary outcome, we used multinomial logistic regression to calculate maximum likelihood estimates of odds ratios (ORs) and 95% confidence intervals (CIs) for course trajectories. The "quick remission" trajectory served as a reference group.

To examine the effect of treatment received during the study, which was not offered to all participants in this naturalistic study, we looked for effect-modification by dichotomizing the group on having received pharmacologic treatment after baseline (including treatment started on baseline itself) or not. We performed stratified analyses on primary outcomes and modeled interaction terms in the regression analyses with total MSM score to estimate significance of effect-modification if present.

We analyzed the effect of both the total MSM score and its components independently. *P* values less than .05 were considered significant.

RESULTS

Descriptive

Out of the total sample of 2,981 NESDA participants, exclusion of controls (n = 652) and patients not meeting the inclusion criterion of having an ongoing episode of depression at baseline resulted in a raw sample of 965 depressed persons. Due to missing data among variables required for MSM scores, our second inclusion criterion narrowed this sample down to 829. Regarding gender distribution, age, and education, this sample was comparable to the raw sample. The third inclusion criterion, regarding the availability of follow-up data, resulted in 643 respondents for analysis. Regarding gender distribution, age, and education, this sample was comparable to the raw sample. Moreover, MSM scores were comparable as well: in the sample of 829 subjects, the mean score was 4.92 (SD = 1.20), while in the final sample (n = 643), the mean score was 4.93 (SD = 1.22). See Supplementary eFigure 1 for flowchart of patient disposition.

Of our sample, mean age was 41 years (SD = 12.2), 428 were female (67%), and 304 (47%) had a first depressive episode (Table 1). A total of 560 subjects (87%) suffered from depression for less than or equal to 12 months prior to baseline. Further, 51 (8%) already had a chronic depressive episode at baseline, ie, had been depressed for > 24 months. Of the subjects, 265 (41%) had a severe depression, and 310 (48%) had not used antidepressants at baseline. The median number of antidepressant drugs used at baseline was 1. Twenty-one patients (3%) had used augmentation

Table 1. Demographic and Clinical Characteristics With Distribution Over Categories of Final Sample (N = 643)

Variable	n (%)	Mean (SD)	Median (IQR)
Age, y		41 ± 12.2	
Female gender	428 (67)		
Sample origin			
General population	43 (7)		
Primary care	228 (36)		
Secondary care	372 (58)		
Education			
Basic	64 (10)		
Intermediate	414 (65)		
High	165 (26)		
Depression type			
MDD first episode	304 (47)		
MDD recurrent episode	339 (53)		
Duration of episode (scoring 1–3)			
Acute (≤ 12 mo)	560 (87)	1.21 ± 0.57	1 (2)
Subacute (13–24 mo)	32 (5)		
Chronic (> 24 mo)	51 (8)		
Symptom severity (at baseline) (scoring 1–5)			
Subsyndromal[a]		3.15 ± 0.81	3 (2)
Mild	168 (26)		
Moderate	210 (33)		
Severe without psychosis	265 (41)		
Severe with psychosis[a]			
Antidepressants used in current episode (scoring 0–5)			
None[b]	310 (48)	0.57 ± 0.60	1 (3)
Level 1: 1–2	302 (47)		
Level 2: 3–4	28 (4)		
Level 3: 5–6	3 (0)		
Level 4: 7–10			
Level 5: > 10			
Augmentation used in current episode (0–1)			
Not used	622 (97)	0.03 ± 0.18	0 (1)
Used	21 (3)		
ECT[a] used in current episode (0–1)			
Not used			
Used			
MSM total		4.93 ± 1.22	5 (6)

[a]This information is not available in the Netherlands Study of Depression and Anxiety database.
[b]This item is not scored in the original Maudsley Staging Method.
Abbreviations: ECT = electroconvulsive therapy, IQR = interquartile range, MDD = major depressive disorder, MSM = Maudsley Staging Method.

Table 2. Prediction of Time Being Depressed (% Time Depressed; Linear Regression Model)[a] and Persistent Depression (Logistic Regression Model)[b]

% Time Depressed	B	95% CI	*P* Value
Univariate models of individual items			
Duration	0.076	0.027 to 0.126	.002
Severity	0.061	0.025 to 0.096	.001
Antidepressants	0.055	0.007 to 0.103	.026
Augmentation	0.096	−0.069 to 0.262	.254
Multivariate model of individual items[c]			
Duration	0.079	0.030 to 0.128	.002
Severity	0.058	0.022 to 0.094	.002
Antidepressants	0.037	−0.011 to 0.086	.130
Augmentation	0.064	−0.100 to 0.229	.442
Final model[d]			
MSM score	0.057	0.034 to 0.081	<.001
Persistent Depression	**OR**	**95% CI**	***P* Value**
Univariate models of individual items			
Duration	1.90	1.41 to 2.57	<.001
Severity	1.31	1.08 to 1.59	.007
Antidepressants	1.36	1.05 to 1.77	.020
Augmentation	1.90	0.78 to 4.64	.161
Multivariate model of individual items[e]			
Duration	1.94	1.43 to 2.62	<.001
Severity	1.30	1.07 to 1.60	.010
Antidepressants	1.25	0.95 to 1.65	.105
Augmentation	1.62	0.65 to 4.05	.303
Final model[f]			
MSM score	1.40	1.23 to 1.60	<.001

[a]Linear regression model: to test for the variable "percentage time depressed" as independent variable.
[b]Binary logistic regression model: Maudsley Staging Method score as a dependent variable and "persistent depression" as independent variable. Both models left uncorrected.
[c]Akaike information criterion (AIC): 590.79.
[d]AIC: 595.93.
[e]AIC: 865.85.
[f]AIC: 866.95.
Abbreviations: MSM = Maudsley Staging Method, OR = odds ratio.

medication at baseline. The mean MSM score was 4.9 (SD = 1.2).

Prediction of Course of Illness During Follow-Up

Regarding our primary outcomes, the MSM significantly predicted "percentage time depressed" ($P < .001$) and was significantly associated with "persistent depression" (≥ 50% of the follow-up) (OR = 1.40; 95% CI, 1.23–1.60; $P < .001$) (Table 2). Participants in this group were, on average, depressed for 89% of the follow-up period. Correction for age and sex did not substantially affect these outcomes (available from the authors on request).

We examined how individual model components predicted "percentage time depressed" and depression during follow-up. Except for augmentation, individual model components in both models univariately predicted a chronic depression during follow-up. In the multivariate model, duration and severity in both models predicted a chronic depression during follow-up. Prediction of the secondary outcome course trajectory showed that each point increase on the MSM significantly predicted a worse course of depression over the following 2 years (Table 3). Correction for age and sex did not substantially affect these outcomes.

Table 3. Prediction of Different Course Trajectories[a]

Course Trajectory	n (%)	OR	95% CI	*P* Value
Quick remission course	265 (41)		Reference	
Decline course, moderate severity	165 (26)	1.30	1.10 to 1.53	.002
Decline course, high severity	69 (11)	1.56	1.25 to 1.95	<.001
Chronic course, moderate severity	93 (15)	1.50	1.22 to 1.83	<.001
Chronic course, high severity	51 (8)	1.46	1.13 to 1.88	.004

[a]Final model: $\chi^2_4 = 28{,}625$, $P < .001$. Multinomial logistic regression model for showing maximum likelihood estimates of odds ratios (OR) and 95% confidence intervals (95% CI) for all courses of depressive symptoms in relation to Maudsley Staging Method scores. Quick remission was taken as reference. Model left uncorrected.

Sensitivity Analyses

Stratification of the predictions for those who received pharmacologic treatment or not showed slightly lower estimates in the "received treatment" group, indicating some modification of effect. However, for the prediction of "percentage time depressed," stratification resulted in absence of significance ($P = .059$) for those who did receive treatment. The MSM was significantly associated with "persistent depression" (≥ 50% of the follow-up) in both the subgroup that received treatment and the subgroup that did not. The interaction MSM × treatment was not significant for any of these outcomes (see Supplementary eTable 2).

The stratified analysis of our secondary outcome revealed an absence of significance for patients who received pharmacologic treatment for the course trajectories "decline course, moderate severity" and "chronic course, high severity." Moreover, patients who had not received pharmacologic treatment showed an absence of significance for the course trajectories "decline course, high severity" and "chronic course, moderate severity" (Supplementary eTable 3). The MSM score by treatment interaction showed no significant results for either course trajectory (Supplementary eTable 4).

DISCUSSION

In the present study, we aimed to assess whether the MSM predicts the 2-year course of MDD in a population-based cohort of depressed subjects. Our study shows that higher MSM scores adequately predict worse depression outcomes in a large and clinically heterogeneous sample of MDD patients recruited in the general population, primary care, and secondary care who were followed up over a 2-year period. Furthermore, this prediction appeared to be independent of treatment provided at baseline or during follow-up. This finding suggests that, in addition to the tertiary population studied by Fekadu et al,[15,19] the MSM can also be used in general psychiatric practices and that the MSM can be used for both prediction of treatment outcome and course of MDD.

In comparison with the sample of Fekadu and colleagues, the current sample has a lower overall MSM score (4.9 [SD = 1.2] vs 10.7 [SD = 2.3]).[15] Indeed, the current sample is more heterogeneous and less often chronically ill, although, in terms of dispersion, our samples appear to have similar variance. In our sample, 8% had a chronic course

at baseline, compared to 61% in the sample of Fekadu et al,[15,19] whereas mild depression was present in 26% versus 10%, respectively. Also, our sample has a greater variety of severity of depression and overall a less extensive treatment history. In the Fekadu sample, 13% had been using only 1 or 2 antidepressants, and most subjects had been using more medications.[15] In our sample, 47% had been using only 1 or 2 antidepressants. To cover the full spectrum of treatment resistance, we also included patients from primary or secondary care who had not been using any antidepressant medication for the current episode at baseline but who did receive treatment during follow-up. By showing no significant interaction (MSM × received treatment), we show that the MSM can predict both course of illness and chances of unfavorable outcome irrespective of treatment during follow-up.

Despite the sample differences between these studies, the MSM performed equally well with regard to predictive validity. First, we found a positive linear correlation between the MSM score and time subjects remained depressed, suggesting that subjects who have a higher MSM score will remain depressed for a longer time. Second, we found that a 1-point increase on the MSM was associated with 1.4-fold increased odds of being depressed for most of the follow-up time. This is comparable to the OR of 1.5 reported in tertiary care.[19] This remarkable similarity suggests that the MSM is applicable in the full spectrum of persons with depression ranging from the general population to tertiary care levels and that it can be validly used for predicting untoward depression outcomes across those different groups.

The individual components of the MSM showed predictive validity. In multivariate analyses, duration and severity contributed significantly to the final models, either linear or logistic, while treatment history no longer did. This could be explained by the fact that severity at baseline correlates with the initiation of pharmacologic treatment (ie, antidepressant use; this correlation was 0.17 [$P < .001$] in our sample).

The difference between how well both models—the multivariate model containing the individual items and the final model containing only the total score—fitted the data was, however, small. As an indication of the optimal fit of these models, we computed the Akaike information criterion (AIC), indicating that explained variance penalized for the number of explanatory variables (smaller is better). The multivariate model fitted slightly less well (AIC: −590.79) than the model with only the MSM score (AIC: −595.93), when tested in a linear regression. When tested in a logistic regression, the reverse was true (AIC: 865.85 for the multivariate model versus AIC: 866.95 for the MSM score only). We therefore propose to retain treatment history in the model. Previous models of quantifying TRD, like the Thase and Rush Staging Method (TRSM)[28] or a variation thereof, the Massachusetts General Hospital Staging method (MGH-S)[29] used only the number of classes of antidepressants (TRSM) or the number of failed trials (MGH-S) to which the patient has not responded. We, however, show that prediction of outcome is improved when clinical variables are included apart from failed treatments.

With regard to our secondary analyses, the MSM significantly predicted chronic course trajectories.[27] These 2-year course trajectories, modeled with accurate information of symptom levels on a per month basis, better represent the course of illness than merely the percentage of time being depressed or a dichotomous distinction between more or less than 50% of time spent in depression. As such, these results confirm the validity of the MSM to predict TRD even further.

A limitation of our study is that NESDA is a naturalistic cohort study, describing the course of depression irrespective of treatment. This potentially limits the scope of our conclusions on treatment resistance. Investigations of treatment effects in naturalistic cohorts like NESDA may be hampered by several factors. These include confounding by indication as a result of physician preferences and current treatment algorithms,[30] meaning that there are reasons for participants to receive different pharmacologic treatments based on their clinical presentation (eg, higher disease severity) and that these reasons then are found to be associated with treatment resistance or other outcomes. Second, power may be insufficient to address all possible treatment strategies. However, most investigations of other tools to predict TRD show that prediction of treatment outcome is possible irrespective of the precise description of the treatment provided.[15,19,28,29] Furthermore, we found little evidence of effect-modification by pharmacologic treatment in our study, so the predictive value of the MSM seemed independent of receiving pharmacologic treatment.

In line with this, another limitation of the NESDA cohort is the limited availability of exact (pharmacologic) treatment data. Although we know the minimal and maximum dose prescribed per antidepressant received and operationalized adequate dosages, we cannot infer the exact time periods of "adequate treatment" (ie, at minimal effective dose for at least 4 weeks) nor compliance to the prescribed treatments. As a result, the number of adequate trials of antidepressants at baseline or the adequacy of received treatment after baseline might have been overestimated.

We used the number of symptoms recorded according to the CIDI to determine severity. Instead, one might expect a more direct score from, for example, the IDS-SR. Here, we followed the initial method proposed by Fekadu et al,[15] which might also better reflect daily clinical practice. This method was chosen to increase the applicability of the MSM for clinical practice. To assess whether our method of scoring severity affected our outcomes, we repeated the main analysis with the IDS score as a severity measure (see Supplementary eTable 5 in Supplementary Results), which did not substantially affect outcomes. An additional analysis in which we left severity out of the MSM and tested a 3-way interaction MSM × severity × received treatment resulted in a nonsignificant finding, both for severity as scored by CIDI criteria ($P = .215$) and for severity as scored by the IDS ($P = .670$). So, our results are not affected by an interaction with severity.

Future studies are needed to establish whether specific treatments are especially effective in certain ranges of the MSM and whether such ranges are sensitive and specific for individual patients. This will be the next step to fully validate the MSM as a profiling tool to guide treatment. Whether additional variables may be helpful to improve this prediction[8] is another issue under debate.[31] The MSM might then be helpful for the apparent clinical need to better predict the course of depression. The MSM might enable clinicians to accurately identify patients who are at risk of developing TRD. An accurate identification could help in offering specific (or more intensified) treatment regimens in an earlier phase than we currently do. Whether this treatment should be another antidepressant, (the addition of) psychotherapy, or other forms of treatment such as neurostimulation remains to be elucidated, but an accurate identification in an earlier phase might provide an important approach to achieve quicker remission of depression. Vice versa, this might also help clinicians to identify patients who have a low risk of an unfavorable course of illness. It should be noted that further study is needed to determine whether patients with lower MSM scores may actually benefit from minimal or only supportive treatment. Until then, it would be advisable to use the MSM in randomized controlled trials to quantify and potentially stratify subjects according to their level of treatment resistance,[32] making it possible to investigate if subjects with different levels of therapy resistance will respond differently to specific treatments.

CONCLUSION

The current study has attempted to validate the predictive value of the MSM as a tool to quantify TRD. With consideration of the sample-related limitations, we conclude that the MSM is a reliable and valid tool to predict poor outcome in depressed patients irrespective of treatment. As an addition to previous work, we show the applicability of MSM in a wider range of primary and secondary care patients with MDD, with varying degrees of prior treatment nonresponse, which is relevant for the description of studied samples in trials investigating TRD. Future aims should be directed to enable the use of MSM scores as a clinically applicable tool to guide clinical treatment selection.

Submitted: January 23, 2017; accepted July 25, 2017.

Published online: January 23, 2018.

Potential conflicts of interest: The authors have no conflict of interest.

Funding/support: This study was funded by University Medical Center Groningen Innovation Fund, project U-11-221, principal investigator Prof R. A. Schoevers, and Fonds NutsOhra, project 1103-068; principal investigator Prof R. A. Schoevers.

Role of the sponsor: The sponsor had no role in the design, analysis, interpretation, or publication of this study.

Additional information: The original data set for the Netherlands Study of Depression and Anxiety is available from http://www.nesda.nl.

Supplementary material: See accompanying pages.

REFERENCES

1. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet*. 2009;373(9665):746–758.
2. Cuijpers P, Berking M, Andersson G, et al. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Can J Psychiatry*. 2013;58(7):376–385.
3. Cuijpers P, Sijbrandij M, Koole SL, et al. The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. *World Psychiatry*. 2013;12(2):137–148.
4. de Maat SM, Dekker J, Schoevers RA, et al. Relative efficacy of psychotherapy and combined therapy in the treatment of depression: a meta-analysis. *Eur Psychiatry*. 2007;22(1):1–8.
5. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163(11):1905–1917.
6. Greden JF. The burden of disease for treatment-resistant depression. *J Clin Psychiatry*. 2001;62(suppl 16):26–31.
7. Ivanova JI, Birnbaum HG, Kidolezi Y, et al. Direct and indirect costs of employees with treatment-resistant and non-treatment-resistant major depressive disorder. *Curr Med Res Opin*. 2010;26(10):2475–2484.
8. Ruhe HG, van Rooijen G, Spijker J, et al. Staging methods for treatment resistant depression: a systematic review. *J Affect Disord*. 2012;137(1–3):35–45.
9. Berlim MT, Turecki G. Definition, assessment, and staging of treatment-resistant refractory major depression: a review of current concepts and methods. *Can J Psychiatry*. 2007;52(1):46–54.
10. Berlim MT, Turecki G. What is the meaning of treatment resistant/refractory major depression (TRD)? a systematic review of current randomized trials. *Eur Neuropsychopharmacol*. 2007;17(11):696–707.
11. 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.
12. Souery D, Papakostas GI, Trivedi MH. Treatment-resistant depression. *J Clin Psychiatry*. 2006;67(suppl 6):16–22.
13. Malhi GS, Parker GB, Crawford J, et al. Treatment-resistant depression: resistant to definition? *Acta Psychiatr Scand*. 2005;112(4):302–309.
14. Ruhe HG, Huyser J, Swinkels JA, et al. Switching antidepressants after a first selective serotonin reuptake inhibitor in major depressive disorder: a systematic review. *J Clin Psychiatry*. 2006;67(12):1836–1855.
15. 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.
16. Spijker J, de Graaf R, Bijl RV, et al. Duration of major depressive episodes in the general population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Br J Psychiatry*. 2002;181(3):208–213.
17. Spijker J, de Graaf R, Bijl RV, et al. Determinants of persistence of major depressive episodes in the general population. results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *J Affect Disord*. 2004;81(3):231–240.
18. Vuorilehto MS, Melartin TK, Isometsa ET. Course and outcome of depressive disorders in primary care: a prospective 18-month study. *Psychol Med*. 2009;39(10):1697–1707.
19. Fekadu A, Wooderson SC, Markopoulou K, et al. The Maudsley Staging Method for treatment-resistant depression: prediction of longer-term outcome and persistence of symptoms. *J Clin Psychiatry*. 2009;70(7):952–957.
20. Penninx BW, Beekman AT, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res*. 2008;17(3):121–140.
21. Wittchen HU. Reliability and validity studies of the WHO—Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res*. 1994;28(1):57–84.
22. Rush AJ, Blacker D, First MB, et al. *Handbook of Psychiatric Measures*. Arlington, VA: American Psychiatric Pub; 2008.
23. Rush AJ, Gullion CM, Basco MR, et al. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477–486.
24. Lyketsos CG, Nestadt G, Cwi J, et al. The Life Chart Interview: a standardized method to describe the course of psychopathology. *Int J Methods Psychiatr Res*. 1994;4(3):143.
25. WHO Collaborating Centre for Drug Statistics Methodology. *Guidelines for ATC classification and DDD assignment 2016*. Oslo, Norway: WHO Collaborating Centre for Drug Statistics Methodology: Norwegian Institute of Public Health; 2016.
26. Penninx BW, Nolen WA, Lamers F, et al. Two-year course of depressive and anxiety disorders: results from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disord*. 2011;133(1–2):76–85.
27. Rhebergen D, Lamers F, Spijker J, et al. Course

trajectories of unipolar depressive disorders identified by latent class growth analysis. *Psychol Med*. 2012;42(7):1383–1396.
28. 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.
29. Fava M. Diagnosis and definition of treatment-resistant depression. *Biol Psychiatry*. 2003;53(8):649–659.
30. Spijker J, Nolen WA. An algorithm for the pharmacological treatment of depression. *Acta Psychiatr Scand*. 2010;121(3):180–189.
31. Peeters FP, Ruhe HG, Wichers M, et al. The Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD): an extension of the Maudsley Staging Method. *J Affect Disord*. 2016;205:365–371.
32. de Kwaasteniet BP, Rive MM, Ruhe HG, et al. Decreased resting-state connectivity between neurocognitive networks in treatment resistant depression. *Front Psychiatry*. 2015;6:28.

Supplementary material follows this article.

THE JOURNAL OF CLINICAL PSYCHIATRY
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL PSYCHOPHARMACOLOGY

Supplementary Material

Article Title: Validity of the Maudsley Staging Method in Predicting Treatment-Resistant Depression Outcome Using the Netherlands Study of Depression and Anxiety

Author(s): Sjoerd M. van Belkum, MD; Hanneke H. Geugies, MSc; Thom S. Lysen, MD; Anthony J. Cleare, PhD; Frenk P. M. L. Peeters, MD, PhD; Brenda W. J. H. Penninx, PhD; Robert A. Schoevers, MD, PhD; and Eric G. Ruhe, MD, PhD

DOI Number: 10.4088/JCP.17m11475

List of Supplementary Material for the article

Disclaimer

This Supplementary Material has been provided by the author(s) as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Supplementary Methods

1. Determining MSM-scores in NESDA

1-1. Duration of current episode before baseline assessment

Duration of 1 year of less was considered acute, between 1 and 2 years was considered sub-acute, and a duration of more than 2 years chronic. For determining duration of episode, the Life chart interview (LCI) at baseline was used. The LCI asked respondents the amount of months in the year before the baseline assessment that were spent with symptoms and the highest perceived burden during these months. Due to difficulties in NESDA to determine the precise length of the depressive episode, episode duration was considered longer than the examined retrospective year if the patient had spent at least 10 months with symptoms and a burden greater than 'not troubled at all' (e.g. not meeting this criterion meant episode duration was considered 'acute').

1-2. Severity

Severity of depression was assessed according to the DSM-IV classification in three categories: (i) mild, (ii) moderate, and (iii) severe. We followed the categorization used by the CIDI [WHO 1998; Wittchen 1994]. Due to exclusion criteria of the NESDA-cohort and lack of information on psychotic symptoms, we could not score for these. Subthreshold depression was not included in the cohorts used for course descriptions and could therefore not be included in the analysis.

1-3. Antidepressants

To assess current treatment failures we made use of treatment counts in NESDA. Respondents were asked to bring their medicine boxes so an inventory of names, dosage

and daily amount could be made, with a specification of medication adherence per drug taken (daily, frequent (>50%), infrequent (<50%), sporadic). Medication use was counted if frequency of use was on a daily basis, if dosage was at least the Daily Defined Dose (DDD) and if it was used for at least 4 weeks (1 month). The DDD is the average daily maintenance dose for use in adults. For the treatment of MDD this is the appropriate dosage for treatment of a moderate to severe depressive episode [WHO 2012]. The MSM specifies the use of the Maudsley Prescribing Guidelines for determining correct daily dose and sets a minimum of at least 6 weeks for adequate use [Fekadu 2009a]. Because no start and stop dates of prescribed drugs were available in NESDA, and uncertainty on when the depressive episode started exactly, medication listed in NESDA is not linked to specific episodes. An extra null category was added to include participants without any previous antidepressant use, for which a score of 0 was appointed.

1-4. Augmentation

The use of augmentation was determined for current medication use and for the whole three-year retrospective period. Medication regarded as augmentation were the following: lithium, anticonvulsants (valproic acid, carbamazepine and lamotrigine), triiodothyronine (T3, synthetic thyroid hormone), pindolol and buspirone. For counting augmentation, the same conditions for frequency, dose and duration applied. Scoring was equal to the proposed scoring in both models.

1-5. ECT

Scores of treatment with electroconvulsive therapy (ECT) could not be determined due to the fact that this was not recorded in the NESDA-database.

eTable 1:

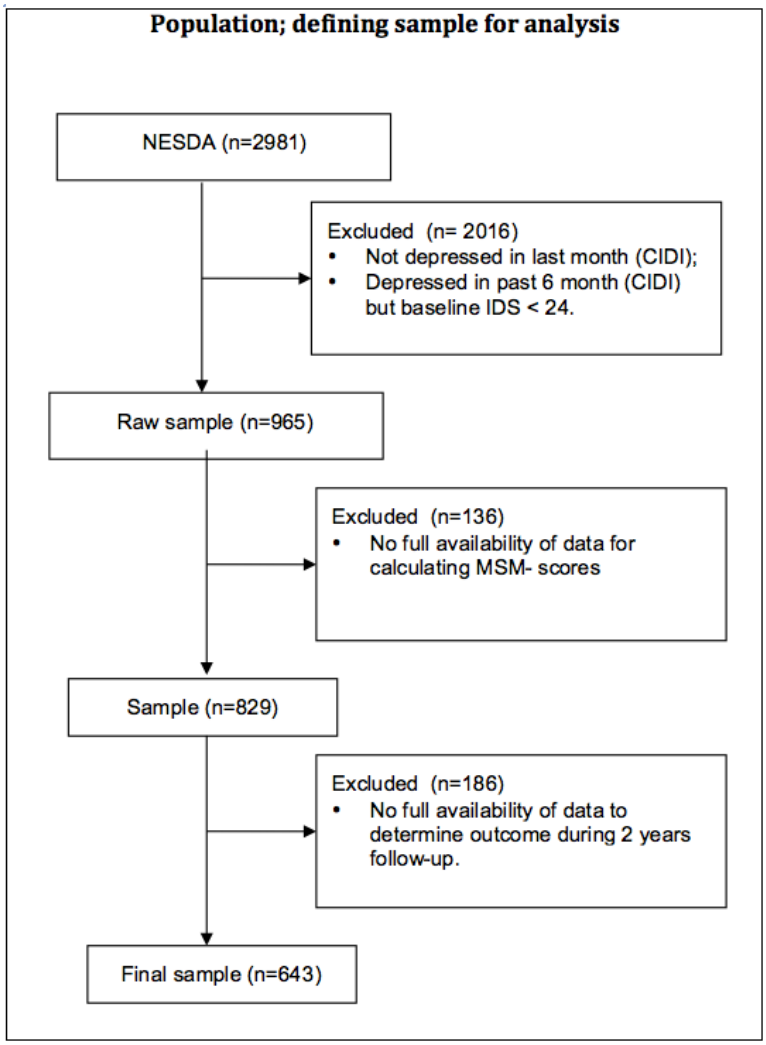
Original MSM-scoring, reprinted with permission (Fekadu 2009, A Multidimensional Tool to Quantify Treatment Resistance in Depression: The Maudsley Staging Method, J. Clin. Psychiatry).

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)

Supplementary Results

eFigure 1:

Flow-chart of patient disposition.



Stratified analyses

eTable 2:

Prediction of time being depressed (A) '% time depressed'; linear regression model) and persistent depression (B); Logistic regression model).

A) % Time depressed	B	95% CI	p-value
MSM stratified by treatment[1]			
MSM –no treatment	0.069	0.024 – 0.125	0.003
MSM –treatment	0.031	-0.001 – 0.064	0.059

B) Persistent depression	OR	95% CI	p-value
MSM stratified by treatment[2]			
MSM –no treatment	1.47	1.11 – 1.94	0.007
MSM –treatment	1.27	1.06 – 1.52	0.020

A) Linear regression model: to test for the variable 'percentage time depressed' as independent variable. B) Binary logistic regression model: MSM score as a dependent variable and the variable 'persistently depressed' as independent variable.

[1]- Interaction MSM*'received treatment' (after baseline): p = 0.191; [2]- Interaction MSM*'received treatment' (after baseline): p = 0.381.

eTable 3:

Prediction of different course trajectories stratified by treatment

	No treatment[1]			Received treatment[2]		
Course trajectory	OR	95% CI	p-value	OR	95% CI	p-value
Quick remission course	Reference			Reference		
Decline course, moderate severity	1.44	1.04 – 2.00	0.026	1.18	0.94 – 1.49	0.155
Decline course, high severity	1.50	0.89 – 2.52	0.128	1.37	1.02 – 1.83	0.035
Chronic course, moderate severity	1.47	0.96 – 2.26	0.080	1.37	1.05 – 1.79	0.022
Chronic course, high severity	1.74	0.90 – 3.35	0.098	1.13	0.82 – 1.55	0.464

[1]- Final model: chi-square (df): 8.616 (4), $p < .071$; [2]- Final model: chi-square (df): 7.676 (4), $p < .104$.

eTable 4:

Prediction of different course trajectories, including the interaction term with received treatment

Course trajectory	OR	95% CI	p-value
Quick remission course		Reference	
Decline course, moderate severity[1]			
MSM-score	1.44	1.04 – 2.00	0.026
Interaction MSM*'received treatment'	0.82	0.55 – 1.22	0.325
Decline course, high severity[2]			
MSM-score	1.50	0.89 – 2.52	0.128
Interaction MSM*'received treatment'	0.91	0.50 – 1.66	0.764
Chronic course, moderate severity[3]			
MSM-score	1.47	0.96 – 2.26	0.080
Interaction MSM*'received treatment'	0.93	0.56 – 1.54	0.780
Chronic course, high severity[4]			
MSM-score	1.74	0.90 – 3.35	0.098
Interaction MSM*'received treatment'	0.65	0.31 – 1.34	0.242

Final model: chi-square (df): 38.546 (12), $p < .001$

eTable 5:

Prediction of time being depressed (A) '% time depressed'; linear regression model) and 'persistent depression' (B); Logistic regression model), using IDS-SR as severity measure, instead of CIDI-methodology (complementary to Table 2).

A) % Time depressed	B	95% CI	p-value
Univariate models of individual items			
Duration	0.076	0.027 – 0.126	0.002
Severity	0.091	0.052 – 0.130	< 0.001
Antidepressants	0.055	0.007 – 0.103	0.026
Augmentation	0.096	-0.069 – 0.262	0.254
Multivariate model of individual items			
Duration	0.060	0.011– 0.109	0.017
Severity	0.077	0.037 – 0.117	< 0.001
Antidepressants	0.029	-0.020 – 0.078	0.251
Augmentation	0.062	-0.102 – 0.226	0.457
Final model			
MSM-score	0.058	0.036 – 0.080	< 0.001

B) Persistent depression	OR	95% CI	p-value
Univariate models of individual items			
Duration	1.90	1.41 – 2.57	< 0.001
Severity	1.66	1.31 – 2.08	< 0.001
Antidepressants	1.36	1.05 – 1.77	0.020
Augmentation	1.90	0.78 – 4.64	0.161
Multivariate model of individual items			
Duration	1.77	1.30 – 2.40	< 0.001
Severity	1.49	1.18 – 1.89	0.001
Antidepressants	1.19	0.90 – 1.57	0.215
Augmentation	1.61	0.64 – 4.05	0.307
Final model			
MSM-score	1.45	1.27 – 1.65	< 0.001

A) Linear regression model: to test for the variable 'percentage time depressed' as independent variable. B) Binary logistic regression model: MSM score as a dependent variable and the variable 'persistent depression' as independent variable. Both models left uncorrected.