Clinical Staging of Major Depressive Disorder:
An Empirical Exploration

Judith Verduijn, MD[a,*]; Yuri Milaneschi, PhD[a]; Albert M. van Hemert, MD, PhD[b]; Robert A. Schoevers, MD, PhD[c]; Ian B. Hickie, MD, FRANZCP[d]; Brenda W. J. H. Penninx, PhD[a]; and Aartjan T. F. Beekman, MD, PhD[a]

ABSTRACT

Objective: Clinical staging has been proposed to supplement psychiatric diagnoses. We examined the construct and predictive validity of a clinical staging model for major depressive disorder (MDD) that distinguishes 8 consecutive stages (0, 1A, 1B, 2, 3A, 3B, 3C, 4) based on symptom severity (Inventory of Depressive Symptomatology [IDS]) and duration (Life Chart Interview) and number of episodes (Composite International Diagnostic Interview [CIDI] based on *DSM-IV* criteria).

Method: This study is a secondary data analysis based on baseline data (collected 2004–2007) and 2-year follow-up assessment (collected 2006–2009) from the Netherlands Study of Depression and Anxiety. 2,333 baseline participants were assigned to the 8 stages of the clinical staging model for MDD, and 2,012 participants were followed up after 2 years. For construct validity, differences between stages in clinical characteristics (eg, severity [IDS], age at onset [CIDI], and comorbid anxiety [CIDI]) were studied. Predictive validity was measured by the extent to which baseline stages predicted 2-year follow-up outcomes (eg, MDD presence [CIDI]).

Results: Later stages scored significantly poorer than early stages on most clinical characteristics and follow-up outcomes (all overall *P* values < .001), confirming validity of the model. This pattern was especially evident in mostly preclinical stages (0, 1A, 1B, 2). Among clinical stages (2, 3A, 3B, 3C, 4), stages characterized by long-lasting symptomatology had at baseline similar and the highest scores (IDS scores: stage 3A = 37.1; stage 4 = 39.0; disability scores: stage 3A = 36.7; stage 4 = 40.6) and compared to preclinical stages had increased probability of having MDD at 2-year follow-up (stage 3A: OR = 4.3; 95% CI, 2.8–6.7; stage 4: OR = 8.1; 95% CI, 5.7–11.5).

Conclusions: This study showed reasonable validity for an MDD staging model that based its stages purely on clinical characteristics. Results suggest that, contrary to the number of episodes, duration of exposure to the depressed state best characterizes clinical (later) stages of MDD. Future studies should test whether modifications to these clinical stages improve the validity of the model.

J Clin Psychiatry 2015;76(9):1200–1208
dx.doi.org/10.4088/JCP.14m09272

[a]Department of Psychiatry and EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands
[b]Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands
[c]University of Groningen, University Medical Center Groningen, Department of Psychiatry, Groningen, The Netherlands
[d]Brain and Mind Institute, University of Sydney, Sydney, Australia
**Corresponding author:* Judith Verduijn, MD, Department of Psychiatry, A.J. Ernststraat 1187, 1081 HL Amsterdam, The Netherlands (j.verduijn@ggzingeest.nl).

Major depressive disorder (MDD) is the second most common cause of years lived with disability worldwide.[1] This is due to its high prevalence,[2,3] relatively early onset,[2] enormous impact on quality of life,[4] and tendency to recur over time.[5] From a public health perspective, it is therefore essential to understand the course of and treatment response in MDD, which are both highly heterogeneous. Studies in first-episode MDD patients showed that their prognosis varied from recovering within 3 months (50%) to remaining depressed for longer than 2 years, often despite treatment (20%),[6] and that 60% developed a recurrent episode.[5]

Clinical staging has been proposed as a valuable tool to better define the highly heterogeneous course and treatment response of a range of psychiatric disorders.[7] According to Gonnella et al, "Staging defines discrete points in the course of individual diseases that are clinically detectable, reflect severity in terms of risk of death or residual impairment, and possess clinical significance for prognosis and choice of therapeutic modality."[8(p637)] Staging has been applied successfully to a variety of physical diseases, especially in cancer, where it has helped to make the case for early detection and to devise specific treatments for specific stages of cancer.[9] It has been hypothesized that application of staging to MDD may have similar benefits.[7,10]

The recently published *DSM-5*[11] appears to have evolved in the direction of clinical staging, incorporating the category *chronic depression* in addition to *first* and *recurrent depression*.[12] However, it is not structured along the lines of clinical staging in which the development of the disorder, from the early at-risk stage to a severe chronic stage, is proposed as the backbone for clinical diagnosis.[7]

For MDD, different staging models have been developed that fit different purposes,[13] such as staging of disease progression[10,14–16] and staging of treatment resistance.[17,18] We focused on a disease-progression staging model that in current epidemiologic work has been the most applied and examined[14] and was derived from the original staging model by McGorry and colleagues developed for psychotic disorders.[7] It divides the course of MDD into 8 stages (Table 1), which are defined primarily by the severity and duration of symptoms and the number of episodes: 3 preclinical stages (stage 0, no depressive complaints but a family history of MDD; stage 1A, mild or nonspecific depressive symptoms; stage 1B, moderate but subthreshold depressive symptoms) and 5 clinical stages (stage 2, first episode of MDD; stage 3A, incomplete remission of first episode of MDD; stage 3B, recurrence or relapse of first episode of MDD; stage 3C, multiple episodes of MDD; stage 4,

severe, persistent, or unremitting MDD). To our knowledge, only 1 study applied the MDD staging model to young persons in the early phases of illness,[19] showing preliminary evidence for construct validity. That study included subjects with a broader range of presenting conditions (MDD, anxiety disorders, psychotic disorders) and did not examine older subjects or significant numbers of patients with more recurrent or persistent disorders.[19]

Our primary aim was to test whether the concept of clinical staging of MDD could be supported by empirical evidence. We applied the MDD staging model to the participants of the Netherlands Study of Depression and Anxiety (NESDA). We tested the extent to which the staging model in its current 8-stage form (1) is associated with a stepwise worsening of a range of clinical indicators (construct validity) and (2) predicts the subsequent 2-year course of MDD (predictive validity).

METHOD

Study Sample

Participants were selected from NESDA, an ongoing, large scale, longitudinal cohort study of the course and determinants of depressive and anxiety disorders.[20] In order to reflect the various developmental stages of depression and anxiety, individuals aged 18–65 years were recruited from the community (19%), general practice (54%), and secondary mental health care (27%). The baseline sample included 2,981 participants (data collected from 2004–2007), consisting of persons with a current or remitted depressive and/or anxiety disorder and healthy controls. Of the 2,981 participants, we selected 2,333 with available baseline data on MDD diagnosis and clinical characteristics. Those excluded (n = 648) were healthy (n = 230), without depressive complaints or a family history of depression; only had a lifetime diagnosis of anxiety (n = 395); or missed information on questionnaires (n = 23).

After 2 years, a face-to-face follow-up assessment was conducted (data collected from 2006–2009). NESDA was approved centrally by the Ethical Review Board of the VU University Medical Centre and subsequently by the local review boards of each participating center, and all participants signed written informed consent.

Assessment of MDD Diagnosis and Clinical Characteristics

Presence of MDD was determined according to *DSM-IV* criteria using the Composite International Diagnostic Interview (CIDI, version 2.1).[21] Number of MDD episodes was derived from the CIDI. Severity of depressive symptoms, 1-week recency, was measured with the 30-item self-report Inventory of Depressive Symptomatology (IDS).[22] Whether participants had first-degree family members with depression was assessed using the family tree method.[23] Duration of the current episode in months and percentage of time spent with depressive symptoms during the past 3 years were calculated using the Life Chart Interview.[24]

Clinical Points

- Clinical staging is a new approach to separate patients with MDD into more homogenous groups. It might direct clinicians to early detection of MDD and the development of specific treatments for each MDD stage.
- Our evaluation of the MDD staging model suggested that the preclinical stages, but not clinical stages, are separable from each other on the basis of clinical characteristics.
- Our evaluation of the MDD staging model suggested that, instead of the number of episodes, duration of exposure to the depressed state best characterizes clinical stages.

Application of the Clinical Staging Model

We used recently proposed criteria on stage-assignment strategies[19] as a starting point to assign participants to the model's 8 stages[14] (Table 1, Figure 1). First, participants were divided into those with or without current (6-month recency) MDD at baseline.

Participants with current MDD were assigned to stages 2, 3A, 3B, 3C, or 4 depending on their type of MDD (first, recurrent), number of episodes (2, ≥ 3), severity of depressive symptoms (≤ mild, ≥ moderate), duration of current episode (1–6, > 6–24, or > 24 months), and percentage of time spent with symptoms in the past 3 years (< 80%, ≥ 80%). Participants were chronically depressed if symptoms were present for more than 24 concatenated months (based on *DSM-5* criteria[11]) or if they had depressive symptoms for ≥ 80% of the time in the past 3 years. Since at least 4 months with few symptoms are necessary to be *recovered*,[25] by using the 80% cutoff, we ensured that participants with a 1- or 2-month interruption in symptoms were still recognized as chronically depressed. Participants experienced a first episode if symptoms lasted ≤ 6 months and were in incomplete remission if symptoms lasted > 6 to ≤ 24 months. The cutoff of 6 months has been a practical choice based on longitudinal naturalistic cohort studies that indicate that the median duration of episodes ranges between 3–6 months.[6,26,27] Participants without current MDD were divided into those without lifetime MDD and those with remitted MDD. For participants who never experienced MDD, assignment to stages 0, 1A, or 1B depended on depression severity score and a positive family history. If participants had remitted MDD, we classified them as stage 1B, assuming they were at high risk for a new MDD episode.

Demographic Characteristics

At baseline, data on gender, age, and years of education were assessed.

Construct Validity

We tested for construct validity by examining whether the more advanced stages of the MDD staging model were associated with poorer scores on clinical MDD characteristics and with other variables that form an indication of disease severity, such as received treatment, comorbidity, and impact on functioning. Although we acknowledge that some of these

Table 1. Description of Each MDD Stage[a] With Criteria Used to Assign Clinical Stages[b]

Stage	Description	Criteria Used to Assign Stage[c]
0	Increased risk of anxiety or depressive disorder; no symptoms currently (target population for recruitment: first-degree teenage relatives of probands)	No MDD or anxiety disorder in lifetime IDS ≤ 13 (normal) First-degree family member with depressive disorder
1A	Mild or nonspecific symptoms of anxiety or depression, including neurocognitive deficits of severe mood disorder; mild functional change or decline	No MDD or anxiety disorder in lifetime IDS 14–25 (mild)
1B	Ultrahigh risk: moderate but subthreshold symptoms of anxiety or depression, with moderate neurocognitive changes and functional decline to caseness (GAF < 70)	Group 1: No MDD or anxiety disorder in lifetime IDS ≥ 26 (≥ moderate) Group 2: MDD or comorbid disorder[d] in lifetime No current MDD
2	First episode of MDD; full-threshold disorder with moderate to severe symptoms, neurocognitive deficits, and functional decline (GAF 30–50)	Current MDD, first episode If MDD present in month prior to interview → short duration If MDD not present in month prior to interview → short/middle duration and IDS ≤ 25 (≤ mild)
3A	Incomplete remission from first episode of care; could be linked or fast-tracked to stage 4	Current MDD, first episode If MDD present in month prior to interview → middle duration If MDD not present in month prior to interview → short/middle duration and IDS ≥ 26 (≥ moderate)
3B	Recurrence or relapse of depressive disorder which stabilizes with treatment at a level of GAF, residual symptoms, or neurocognition below the best level achieved following remission from first episode	Group 1: Current MDD, second episode Short/middle duration Group 2: Current MDD, recurrent but unknown number of episodes Short/middle duration
3C	Multiple relapses, provided worsening in clinical extent and impact of illness is objectively present	Current MDD, multiple episodes Short/middle duration
4	Severe, persistent, OR unremitting illness as judged on symptoms, neurocognition, and disability criteria; note: could fast track to this stage at first presentation through specific clinical and functional criteria (from stage 2) or alternatively by failure to respond to treatment (from stage 3A)	Current MDD, first or recurrent episode Long duration

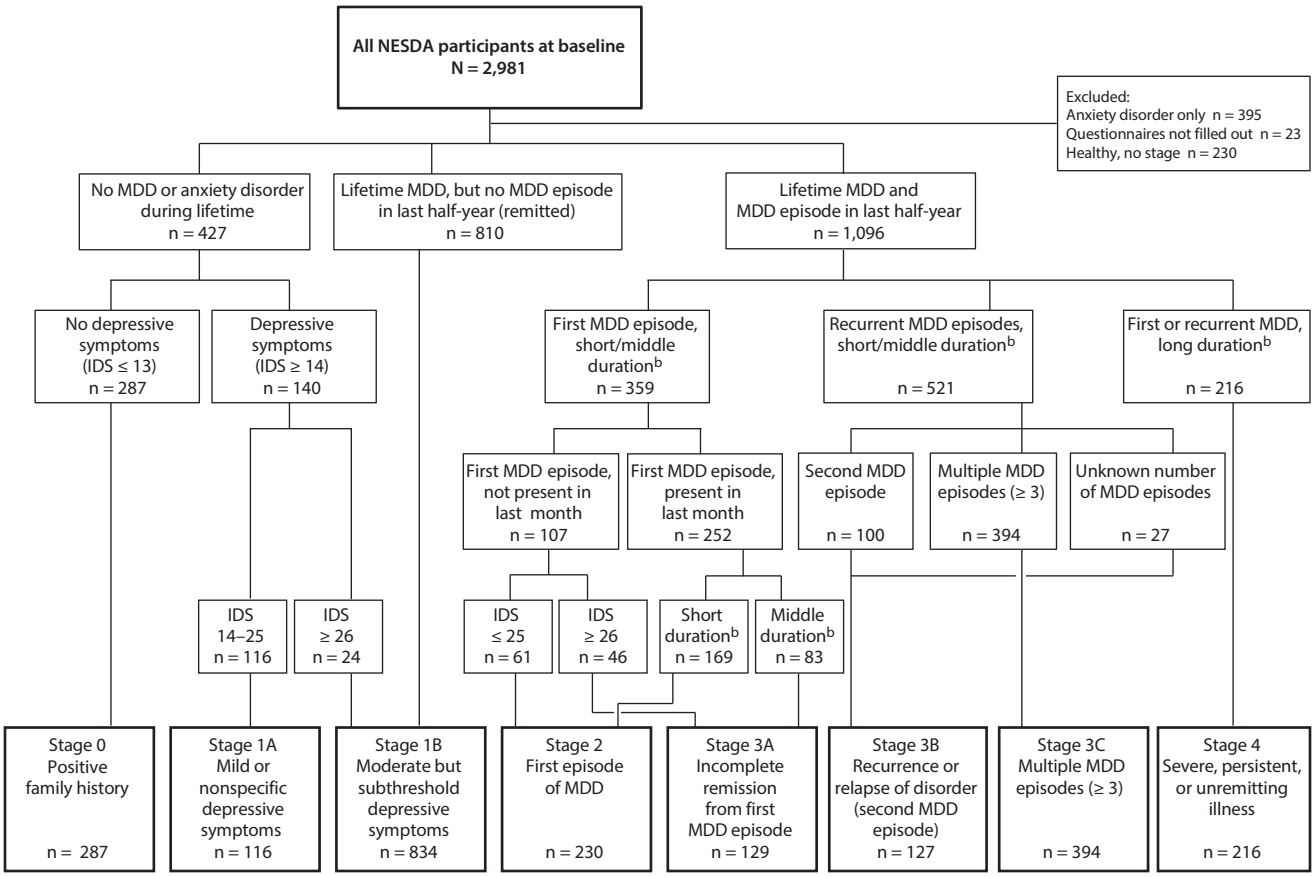
[a]Adapted from Hetrick et al,[14] with permission.
[b]Based on Hickie et al.[19]
[c]Short duration indicates symptoms continuously present ≤ 6 months and present < 80% of the time in the previous 3 years. Middle duration indicates symptoms continuously present > 6 to ≤ 24 months and present < 80% of the time in the previous 3 years. Long duration indicates symptoms continuously present > 24 months and/or present ≥ 80% of the time in the previous 3 years. Current MDD is defined as 6 months' recency. IDS severity categories from http://www.ids-qids.org/index2.html#table2; accessed November 26, 2013.
[d]Comorbid disorder = MDD + anxiety disorder.
Abbreviations: GAF = Global Assessment of Functioning, IDS = Inventory of Depressive Symptomatology, MDD = major depressive disorder.
Symbol: → = then.

characteristics were also used to assign participants to some stages of the model (eg, depression severity for stages 0–3A), describing these characteristics across all stages provided a more complete insight into the overall construct validity of the staging model. Number of experienced MDD symptoms during one's lifetime and age at onset were assessed with the CIDI. Information on lifetime suicide attempts and suicidal ideation in the past week was collected using the Beck Scale for Suicide Ideation (BSI).[28] Participants were considered to be under treatment if they used an antidepressant (1-month recency) or received psychotherapy (6-month recency). Medication use was verified by drug container inspection and registered according to the Anatomical Therapeutic Chemical (ATC) classification.[29] Antidepressants included inter alia selective serotonin reuptake inhibitors (ATC classification: N06AB) and tricyclic antidepressants (ATC classification: N06AA). Information on received psychological treatment (formal psychotherapy or counseling) was acquired by self-report. Presence of any current (6-month recency) comorbid diagnosis of alcohol disorder (dependence or abuse), dysthymia, or anxiety disorder (panic disorder, social phobia, agoraphobia, generalized anxiety disorder) was established using the CIDI. Severity of anxiety symptoms was assessed with the 21-item self-report Beck Anxiety Inventory (BAI).[30] Disability level was measured using the World Health Organization Disability Assessment Schedule 2.0 (WHODAS-II),[31] with score range of 0%–100% (no to full disability).

Predictive Validity

The ability of the model to predict 2-year follow-up clinical outcomes was examined for 3 measures: (outcome A) presence of a current (6-month recency) MDD diagnosis at follow-up; (outcome B) duration of depressive symptoms during the follow-up period, measured with the Life Chart Interview as the percentage of the follow-up period spent

Figure 1. Assignment of Participants to the Clinical Stages of MDD[a]



[a]Participants at the NESDA baseline interview are assigned to the 8 stages of the clinical staging model for MDD.
[b]Short duration indicates symptoms continuously present ≤ 6 months and present < 80% of the time in the previous 3 years. Middle duration indicates symptoms continuously present > 6 to ≤ 24 months and present < 80% of the time in the previous 3 years. Long duration indicates symptoms continuously present > 24 months and/or present ≥ 80% of the time in the previous 3 years.
Abbreviations: IDS = Inventory of Depressive Symptomatology, MDD = major depressive disorder, NESDA = Netherlands Study of Depression and Anxiety.

with depressive symptoms; and (outcome C) disability level at follow-up, assessed with the WHODAS-II.

Statistical Analyses

Variables are reported as percentages, means ± SD, or medians + interquartile range, as appropriate. Construct validity was tested by examining the differences in validator variable scores across stages using analyses of variance (ANOVAs), Kruskal-Wallis tests, and χ^2 statistics. For continuous measures, we performed post hoc tests to examine if consecutive stages differed significantly from each other (eg, stage 1B vs 1A) using Games-Howell procedure after ANOVA (since equal group size and/or equal variance could not be assumed) and multiple Mann-Whitney tests with Bonferroni correction after Kruskal-Wallis tests.

Predictive validity was tested by examining the differences across stages in 2-year follow-up outcomes. For the dichotomous outcome, MDD at follow-up, differences across stages were examined using χ^2 statistics. Effect sizes were odds ratios from logistic regression. As measures of explained variance, both Nagelkerke pseudo R^2 from logistic regression and the *C* statistic from area-under-the-curve analysis were used. For the continuous outcome disability at follow-up, differences across stages were examined using ANOVA and the appropriate post hoc test. The R^2 value from linear regression was interpreted as a measure of explained variance. Because of its nonnormal distribution, for the continuous outcome duration of symptoms during follow-up, differences across stages were examined using Kruskal-Wallis tests followed by post hoc tests. For both continuous outcomes, effect sizes were Hedges *g*.[32] All analyses were repeated with adjustments for age and gender and were conducted using SPSS version 20.0.[33] Significance level was set at $P < .05$, 2-tailed.

RESULTS

Demographic Characteristics

Table 2 shows the number of participants in each stage and their demographics. The 2,333 participants had a mean age of 41.7 (SD = 12.9) years, and 67.9% were female. While the gender proportion did not differ across stages, the mean age and years of education differed significantly without a clear pattern.

Table 2. Distribution of Demographic Characteristics and Construct Validators Across Clinical Stages

Variable	Stages (total N = 2,333)							Across Stages		
	0 (n = 287)	1A (n = 116)	1B (n = 834)	2 (n = 230)	3A (n = 129)	3B (n = 127)	3C (n = 394)	4 (n = 216)	Statistic	*df*	*P* Value
Demographic characteristics											
Gender, % female	62.4	67.2	70.3	64.8	70.5	69.3	69.8	63.4	$\chi^2 = 10.4$	7	.169
Age, mean (SD), y	38.2 (15.3)	43.4* (13.9)	43.6 (12.5)	39.3* (12.2)	38.2 (12.5)	42.4 (12.1)	41.1 (12.0)	42.7 (11.7)	$F = 9.0$	7	<.001
Education, mean (SD), y	12.9 (3.2)	11.9 (3.2)	12.4 (3.2)	11.6* (3.1)	10.9 (3.1)	11.2 (3.2)	12.3* (3.2)	11.1* (3.3)	$F = 11.8$	7	<.001
Construct validators											
Age at onset, mean (SD), y	NA	NA	28.4 (12.1)	29.0 (12.8)	29.9 (12.5)	31.3 (12.4)	24.4* (11.6)	26.0 (12.4)	$F = 10.5$	5	<.001
BAI score, mean (SD)	2.8 (3.0)	8.8* (5.0)	10.5* (8.6)	16.5* (11.8)	22.0* (10.7)	15.1* (9.0)	16.4 (10.5)	22.0* (11.6)	$F = 123.3$	7	<.001
Duration, mean (SD), %[a]	1.2 (5.0)	8.0* (17.5)	10.8 (18.7)	27.7* (19.1)	43.1* (20.4)	24.7* (18.6)	28.1 (19.1)	95.5* (7.2)	$F = 736.0$	7	<.001
Disability, mean (SD), %	6.1 (8.0)	16.0* (8.7)	18.0 (14.3)	29.6* (15.8)	36.7* (13.0)	29.6* (15.9)	31.1 (15.6)	40.6* (17.4)	$F = 165.5$	7	<.001
No. of MDD symptoms, median (IQR)[b]	0 (0.0–3.0)	2* (0.0–5.0)	7* (6.0–8.0)	8* (7.0–9.0)	8 (7.0–9.0)	8 (7.0–9.0)	8 (7.0–9.0)	8 (7.8–9.0)	$\chi^2 = 935.4$	7	<.001
No. of MDEs, median (IQR)[c]	NA	NA	1 (1.0–3.0)	1* (0)	1 (0)	2* (0)	6* (4.0–12.0)	1* (1.0–1.0)	$\chi^2 = 937.3$	5	<.001
Family history of depression, %	100	64.7	73.6	78.6	69.0	74.8	80.5	79.9	$\chi^2 = 112.6$	7	<.001
IDS score, mean (SD)	5.8 (3.7)	18.6* (3.5)	18.3 (10.3)	29.5* (12.8)	37.1* (8.7)	29.6* (11.9)	30.5 (11.9)	39.0* (11.0)	$F = 305.4$	7	<.001
Suicide ever attempt, %	0.7	2.7	10.9	20.4	29.7	13.4	17.8	27.8	$\chi^2 = 132.6$	7	<.001
Suicidal ideation, %	1.4	0.9	5.2	17.8	24.8	15.1	22.3	42.1	$\chi^2 = 294.3$	7	<.001
Treatment, %											
Antidepressant use	0.7	4.3	21.9	44.5	51.6	44.8	35.1	55.8	$\chi^2 = 323.3$	7	<.001
Psychotherapy	5.2	9.5	20.6	50.9	62.0	48.0	46.2	50.9	$\chi^2 = 348.9$	7	<.001
Single treatment	5.2	10.3	26.6	48.0	57.8	32.8	37.4	44.7	$\chi^2 = 580.8$	14	<.001
Both treatments	0.3	1.7	7.9	23.6	28.1	29.6	22.1	31.2			
Comorbid anxiety, %	0.0	0.0	35.6	62.2	68.2	56.7	63.7	75.9	$\chi^2 = 562.9$	7	<.001
Comorbid dysthymia, %	0.0	3.4	2.0	22.2	21.7	10.2	6.6	64.8	$\chi^2 = 743.6$	7	<.001
Comorbid alcoholism, %	7.3	7.8	8.2	13.0	7.0	11.0	13.5	13.0	$\chi^2 = 17.3$	7	.015

[a]Months spent with depressive symptoms in past 3 years, displayed as a percentage.
[b]Number of MDD symptoms experienced during lifetime.
[c]Number of MDD episodes experienced during lifetime.
*Significantly different compared to previous stage (eg, stage 1B vs stage 1A); significance level is .05 for post hoc analysis of variance and .007 for post hoc Kruskal-Wallis.
Abbreviations: BAI = Beck Anxiety Inventory, IDS = Inventory of Depressive Symptomatology, IQR = interquartile range, MDD = major depressive disorder, MDE = major depressive episode, NA = not applicable.

At 2-year follow-up, 2,012 participants (86.2%) were reinterviewed. Compared to responders, nonresponders did not differ in gender or age and were more likely to be classified in one of the clinical stages (3A, 4) with long-lasting symptoms.

Construct Validity

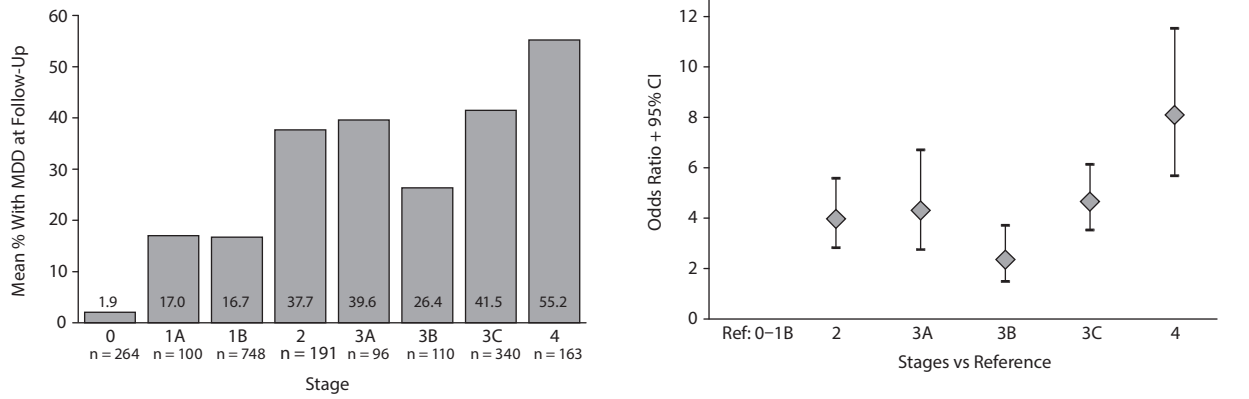
All clinical characteristics examined for construct validity showed significant differences across stages (Table 2). For most validators, a significant pattern of poorer scores moving from early to later stages (referred to as an upward linear pattern) was recognized, which was mainly explained by poorer clinical characteristics in the highest stage (stage 4) than in the lowest stage (stage 0). A common finding of the post hoc analyses was that the upward linear pattern was discontinued around stages 3A, 3B, or 3C, with stage 3B having a generally (significantly) more favorable score than stage 3A. For instance, post hoc analyses for depression severity, anxiety severity, duration, and disability all showed an upward worsening over stages 0 to 3A but then a drop in severity, duration, and disability scores between stages 3A and 3B, followed by slightly poorer characteristics in stage 4. When examined in more detail, the depression severity scores of stages 2, 3B, and 3C (stages that differ only in number of episodes) were almost identical and not significantly different from each other. A similar pattern was found for stages with long-lasting symptoms (3A, 4). Furthermore, the first group of stages (2, 3B, 3C) had significantly more favorable scores on most validators compared to the second group of stages (3A, 4). Of the stages with current MDD (stages 2–4), stages 2, 3A, and 3B had a rather similar age at onset, which was significantly older than the age at onset in stages 3C and 4. When compared to those in stage 3C, those in stage 4 experienced significantly fewer MDD episodes with significantly longer duration of symptoms. Both comorbid anxiety and dysthymia showed a similar pattern as the severity, duration, and disability scores, with the highest proportion of comorbid cases in stage 4.

Predictive Validity

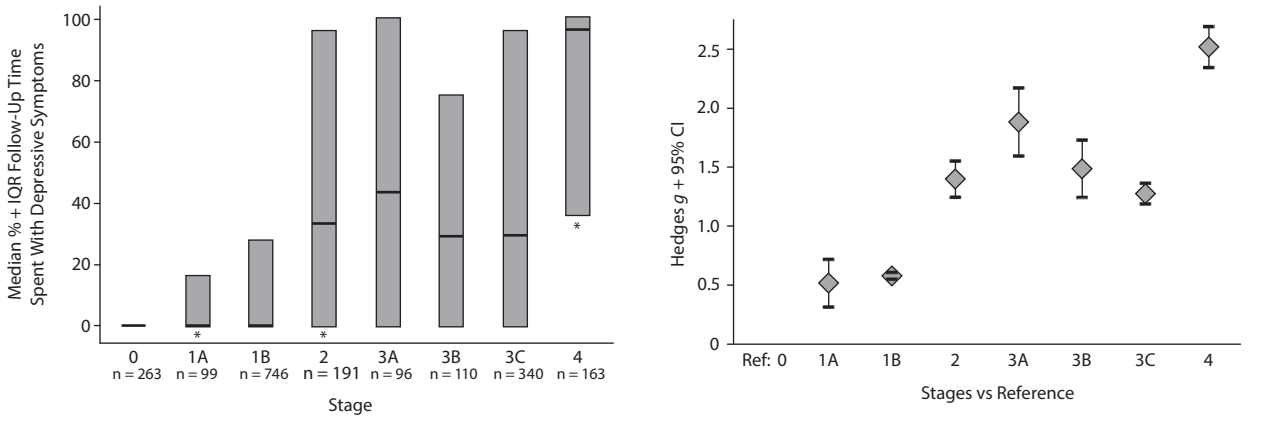
Figure 2 shows the results of predictive validity analyses. Presence of a current MDD diagnosis at 2-year follow-up was

Figure 2. Association Between Clinical Stages and Outcomes at 2-Year Follow-Up

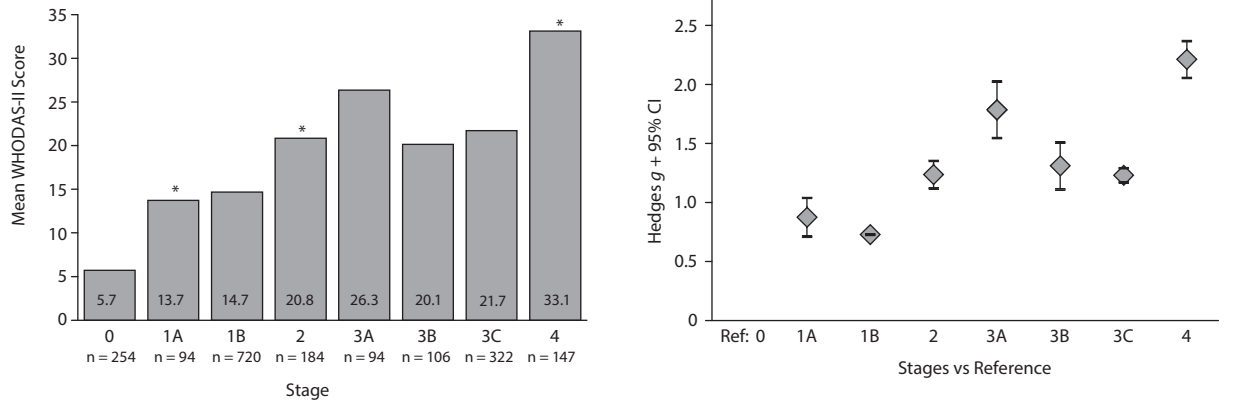
A. Percentage of Participants per Stage With Current MDD at 2-Year Follow-Up[a,b]



B. Median Duration of Depressive Symptoms Over 2-Year Follow-Up per Stage[c,d]



C. Mean Disability (WHODAS-II) Score per Stage at 2-Year Follow-Up[e,f]



[a]Difference across stages in percentage of participants with MDD at 2-year follow-up was examined with the χ^2 test. Effect sizes are odds ratios with 95% confidence intervals derived from logistic regression. Figure is based on the unadjusted analyses.
[b]$\chi^2_7 = 256.8$, $P < .001$.
[c]Difference across stages in median time spent with depressive symptoms over 2-year follow-up was examined with Kruskal-Wallis tests, and comparisons between consecutive stages were performed with Mann-Whitney tests with Bonferroni correction. Effect sizes were calculated as Hedges *g* with a 95% confidence interval. Actual number of participants in stages 0, 1A, and 1B is lower since some participants did not fill out the questionnaire. Figure is based on the unadjusted analyses.
[d]$\chi^2_7 = 472.9$, $P < .001$.
[e]Difference between mean disability outcome at 2-year follow-up was examined with ANOVA, and comparisons between consecutive stages were performed with post hoc analyses: Games-Howell (since equal group size and/or equal variance could not be assumed). Effect sizes were calculated as Hedges *g* with 95% confidence interval. Actual number of participants in each stage is lower since some participants did not fill out the questionnaire. Figure is based on the unadjusted analyses.
[f]$F_7 = 64.7$, $P < .001$.
*Significantly different compared to previous stage; significance level is .05 for post hoc ANOVA and .007 for post hoc Kruskal-Wallis.
Abbreviations: ANOVA = analysis of variance, CI = confidence interval, IQR = interquartile range, MDD = major depressive disorder, WHODAS-II = World Health Organization Disability Assessment Schedule 2.0.

found in 517 of the participants (25.7%). Analyses showed a pattern of increasing MDD cases across stages 0, 1A, 1B, 2, and 3A and a substantial drop at stage 3B. To perform logistic regression, all preclinical stages (0, 1A, 1B) were combined to form the reference group because of the small number of incident MDD episodes in the lowest stage (stage 0). Unadjusted logistic regression showed an explained variance of 15.9% with a *C* statistic of 0.70 (95% CI, 0.68–0.73).

The median percentage of time spent with depressive symptoms during follow-up was 8.3% (interquartile range [IQR], 0%–58.3%). An upward linear pattern was found across stages, with stage-4 patients spending 96% of the follow-up period (IQR, 32.0%–100%) with depressive symptoms. Significantly more time was spent with depressive symptoms between consecutive stages 0–1A and 1B–2; however, this upward line was not continued in clinical stages 2, 3A, 3B, and 3C. Effect sizes were large when comparing stage 0 with stages 2 to 4 (1.28–2.52) but not when comparing consecutive stages (0.13–0.66).

The mean disability score at follow-up was 17.5 (SD = 16.0). Again, an upward linear pattern was observed with a substantial drop at stage 3B. Unadjusted analysis showed an explained variance of 19.1%. Effect sizes were large when comparing stage 0 with all other stages (0.73–2.21) and between stages 0–1A only when consecutive stages were compared (0.07–0.87).

All analyses were repeated with adjustment for age and gender; no substantial change in predictive outcomes was found (data not shown).

DISCUSSION

Our study aim was to test whether the concept of clinical staging of MDD could be supported by empirical evidence. For this purpose, we tested construct and predictive validity of the staging model for MDD.[7,14] Our results showed that the staging model has reasonable validity across stages. Important clinical characteristics, such as severity, duration, and disability, scored progressively worse over stages 0 to 4, suggesting construct validity. Encouraging predictive validity for clinical staging was suggested by the model's performance in predicting 2-year follow-up outcomes (eg, presence of MDD [*C* statistic = 0.70]). However, when results were examined in more detail, we found that, while major clinical characteristic scores and follow-up outcomes were steadily poorer in mostly preclinical stages (0, 1A, 1B, 2), no robust differences were detected between consecutive clinical stages (2, 3A, 3B, 3C, 4). In addition, long-lasting symptom stages (3A, 4) reported poorer characteristics as compared to recurrent depression stages (3B, 3C). Moreover, stages that only differed in number of episodes (2, 3B, 3C) had rather similar clinical characteristic scores and follow-up outcomes.

Our findings seem to suggest that staging MDD based on number of episodes may be less robust compared to staging MDD based on illness duration, when the aim is to assess disease progression. The rationale for considering the number of episodes as an index of disease progression is consistent with the kindling hypothesis,[34] which predicts that prior MDD episodes damage the brain, leading to increased susceptibility for further episodes that become less contingent on environmental stress. Consequently, recurrence rates are expected to be higher in patients who have experienced multiple episodes. Consistent with this hypothesis, studies have shown that the relationship between stress and MDD onset decreases when patients experience more MDD episodes[35] and that patients who experienced multiple episodes had more distinct brain alterations[36,37] and a higher risk of relapse/recurrence.[38] Other studies had contrasting results. One study showed that response to antidepressant treatment in patients with an acute depressive episode was not associated with the number of experienced episodes[39]; other studies showed that brain alterations are associated not only with number of relapses but also with longer cumulative duration of illness[36,40]; and, moreover, studies showed that patients who had residual (long-lasting) symptoms after response to acute treatment were at higher risk for a severe course.[41,42] These findings suggest that disease progression may be better indexed by symptom duration rather than by number of episodes. This latter concept is better supported by our findings whereby stages with long-lasting symptoms (3A, 4) had the poorest clinical characteristic scores and follow-up outcomes. Moreover, a recent review showed that treatment directed at residual symptoms prevented MDD relapses/recurrence in patients with MDD.[43] This line of evidence emphasizes the need to test whether staging models should be constructed based on duration and residual symptoms, at least for the overt phases of illness,[15,44] to improve the model's validity. Further, when staging is studied in more detail, we should strive for the most concise staging model for use in clinical practice. Our results might suggest a simplified method of staging, with stages based on duration of symptoms. However, the current study aimed to test the staging model as developed by Hetrick et al,[14] and, therefore, we used the full model.

Another avenue to improve staging of MDD may be extension of the current construction of stages with profilers. Profilers can be psychological, neuropsychological, neuroimaging, or other neurobiological markers that predict the onset of MDD and/or the progression rate/route over stages. Profilers that might predict progression from attenuated syndromes (stages 1A–1B) to established disorders (stages 2+) include lower melatonin levels[45] and microstructural white matter changes.[46] NESDA studies suggest that comorbid anxiety[47] and childhood trauma[48] might be profilers for MDD course. Whether adding specific profilers improves the validity of the staging model remains to be tested.

The main strengths of our study are its large number of well-characterized patients across the whole adult age range representing different developmental stages of MDD and its longitudinal design. However, our study also has some limitations. First, the degree of fit between the model developed here and those used currently in clinical settings can only be approximated. Importantly, our study included

subjects who were older and more likely to have recurrent or persistent disorders as compared to those in the clinical studies previously conducted.[19,49] Furthermore, in our study, we assigned persons to a stage regardless of their treatment status, whereas in a clinical setting, progression to a higher stage is based, among other things, on the response to received treatment.[19] Our stages 2 through 4, therefore, included a more heterogeneous group of patients (treated: yes/no, treatment response: yes/no); this reduced the differences between stages and might have led to an underestimation of the construct and predictive validity. Moreover, although treatment might be of influence on the course of depression, we decided not to include treatment as a covariate to our predictive analyses since a previous NESDA study into the 2-year course of depression showed in a multivariate model that antidepressant use is not a predictor of course.[47] Other studies should examine the effect of treatment response on stage assignment and progression. Second, some of the characteristics used to test the construct validity were also used to assign participants to some stages (depression severity, number of episodes). However, this is not a major issue since describing these characteristics across all stages provided a more complete insight into the overall construct validity of the staging model, and, more importantly, these characteristics (depression severity, number of episodes) showed identical construct validity compared to characteristics not used for stage assigning (anxiety severity, disability). Moreover, the characteristics used for both stage assignment and construct validity showed similar results as the predictive validity analyses, which were independent of stage assignment. Finally, during follow-up, there was selective loss of participants from the long-lasting stages (3A, 4), which could have influenced the predictive validity. However, since those who are lost to follow-up are generally the worst affected cases (worst affected within stages 3A and 4), selective loss tends to decrease the strength of the associations investigated and suggests that the prospective relation found might be even stronger.

Although not a limitation, another important topic of discussion is the frequent co-occurrence of depression and anxiety disorders in the clinical stages of MDD[3,50] and the frequent development of anxiety symptoms early in the course for those that later develop depression (preclinical stages).[16,51] When anxiety is coexistent, the impact is important, both for effects on the patient (eg, for quality of life) and in terms of a worse prognosis. In the current study, we deliberately aimed to test the staging model for depression as currently developed, and, therefore, adjusting our analyses for anxiety was beyond the scope of our study.

In conclusion, this article is one of the first attempts to empirically validate a clinical staging model for MDD. The present findings are encouraging, especially for the mostly preclinical stages (0, 1A, 1B, 2). However, in the clinical stages (2, 3A, 3B, 3C, 4), validity analyses showed no differences between each consecutive stage. The major reason for this is that the number of episodes appears to have poor discriminatory power compared with the duration of exposure to the depressed state (duration, residual symptoms, and, perhaps, severity). Incorporating profilers may further increase the clinical validity of staging models for MDD, ultimately leading to the integration of the model in the *DSM*, which seems to be moving in the direction of clinical staging.

Submitted: May 27, 2014; accepted October 6, 2014.

Potential conflicts of interest: **Dr Beekman** has received grant/research support from Eli Lilly, Janssen, AstraZeneca, and Shire and has been a speaker for Lundbeck and Eli Lilly. The other authors report no potential conflict of interest.

Funding/support: The infrastructure for the NESDA study (www.nesda.nl/nl) is funded through the Geestkracht program of the Netherlands Organization for Health Research and Development (ZonMw, grant number 10-000-1002) and is supported by participating universities and mental health care organizations (VU University Medical Centre, GGZ inGeest, Arkin, Leiden University Medical Centre, GGZ Rivierduinen, University Medical Centre Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Institute for Quality of Health Care [IQ Healthcare], Netherlands Institute for Health Services Research [NIVEL], and Netherlands Institute of Mental Health and Addiction [Trimbos]).

Role of the sponsor: The sponsors of NESDA did not have any influence on the performed analyses or the text as written.

Additional information: The NESDA database is owned by the NESDA consortium. For information on accessing the database, contact Brenda W. J. H. Penninx (b.penninx@vumc.nl).

REFERENCES

1. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2163–2196.
2. Andrade L, Caraveo-Anduaga JJ, Berglund P, et al. The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *Int J Methods Psychiatr Res*. 2003;12(1):3–21.
3. Kessler RC, Berglund P, Demler O, et al; National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003;289(23):3095–3105.
4. Bijl RV, Ravelli A. Current and residual functional disability associated with psychopathology: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Psychol Med*. 2000;30(3):657–668.
5. Solomon DA, Keller MB, Leon AC, et al. Multiple recurrences of major depressive disorder. *Am J Psychiatry*. 2000;157(2):229–233.
6. 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.
7. McGorry PD, Hickie IB, Yung AR, et al. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Aust N Z J Psychiatry*. 2006;40(8):616–622.
8. Gonnella JS, Hornbrook MC, Louis DZ. Staging of disease: a case-mix measurement. *JAMA*. 1984;251(5):637–644.
9. Sobin LH. TNM: evolution and relation to other prognostic factors. *Semin Surg Oncol*. 2003;21(1):3–7.
10. Fava GA, Kellner R. Staging: a neglected dimension in psychiatric classification. *Acta Psychiatr Scand*. 1993;87(4):225–230.
11. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition. Arlington, VA: American Psychiatric Publishing; 2013.
12. American Psychiatric Publishing AdoAPAA. *Highlights of Changes from DSM-IV-TR to DSM-5*. Washington, DC: American Psychiatric Association; 2013:1–19.
13. Fava GA, Rafanelli C, Tomba E. The clinical process in psychiatry: a clinimetric approach. *J Clin Psychiatry*. 2012;73(2):177–184.
14. Hetrick SE, Parker AG, Hickie IB, et al. Early identification and intervention in depressive disorders: towards a clinical staging model. *Psychother Psychosom*. 2008;77(5):263–270.
15. Cosci F, Fava GA. Staging of mental disorders: systematic review. *Psychother Psychosom*. 2013;82(1):20–34.
16. Fava GA, Tossani E. Prodromal stage of major depression. *Early Interv Psychiatry*. 2007;1(1):9–18.
17. 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.

18. Ruhé 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.
19. Hickie IB, Scott EM, Hermens DF, et al. Applying clinical staging to young people who present for mental health care. *Early Interv Psychiatry*. 2013;7(1):31–43.
20. Penninx BW, Beekman AT, Smit JH, et a; NESDA Research Consortium. 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, Gullion CM, Basco MR, et al. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477–486.
23. Fyer AJ, Weissman MM. Genetic linkage study of panic: clinical methodology and description of pedigrees. *Am J Med Genet*. 1999;88(2):173–181.
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;4:143–155.
25. Rush AJ, Kraemer HC, Sackeim HA, et al; ACNP Task Force. Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology*. 2006;31(9):1841–1853.
26. 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(10):809–816.
27. Ormel J, Oldehinkel T, Brilman E, et al. Outcome of depression and anxiety in primary care: a three-wave 3½-year study of psychopathology and disability. *Arch Gen Psychiatry*. 1993;50(10):759–766.
28. Beck AT, Kovacs M, Weissman A. Assessment of suicidal intention: the Scale for Suicide Ideation. *J Consult Clin Psychol*. 1979;47(2):343–352.
29. World Health Organization Collaborating Centre for Drug Statistics Methodology. Anatomical Therapeutic Chemical (ATC) classification. 2010. http://www.whocc.no/atc_ddd_index. Accessed June 23, 2015.
30. Beck AT, Epstein N, Brown G, et al. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*. 1988;56(6):893–897.
31. Chwastiak LA, Von Korff M. Disability in depression and back pain: evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *J Clin Epidemiol*. 2003;56(6):507–514.
32. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *J Educ Behav Stat. 1981;* 6(2):106–128.
33. Corp IBM. *IBM SPSS Statistics for Windows, Version 20.0*. Armonk, NY: IBM Corp.; 2011.
34. Post RM. Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *Am J Psychiatry*. 1992;149(8):999–1010.
35. Kendler KS, Thornton LM, Gardner CO. Stressful life events and previous episodes in the etiology of major depression in women: an evaluation of the "kindling" hypothesis. *Am J Psychiatry*. 2000;157(8):1243–1251.
36. Lorenzetti V, Allen NB, Fornito A, et al. Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. *J Affect Disord*. 2009;117(1–2):1–17.
37. McKinnon MC, Yucel K, Nazarov A, et al. A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. *J Psychiatry Neurosci*. 2009;34(1):41–54.
38. Bockting CL, Spinhoven P, Koeter MW, et al; Depression Evaluation Longitudinal Therapy Assessment Study Group. Prediction of recurrence in recurrent depression and the influence of consecutive episodes on vulnerability for depression: a 2-year prospective study. *J Clin Psychiatry*. 2006;67(5):747–755.
39. Dodd S, Berk M, Kelin K, et al. Treatment response for acute depression is not associated with number of previous episodes: lack of evidence for a clinical staging model for major depressive disorder. *J Affect Disord*. 2013;150(2):344–349.
40. Sheline YI, Sanghavi M, Mintun MA, et al. Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *J Neurosci*. 1999;19(12):5034–5043.
41. Judd LL, Akiskal HS, Maser JD, et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Arch Gen Psychiatry*. 1998;55(8):694–700.
42. Zajecka J, Kornstein SG, Blier P. Residual symptoms in major depressive disorder: prevalence, effects, and management. *J Clin Psychiatry*. 2013;74(4):407–414.
43. Guidi J, Fava GA, Fava M, et al. Efficacy of the sequential integration of psychotherapy and pharmacotherapy in major depressive disorder: a preliminary meta-analysis. *Psychol Med*. 2011;41(2):321–331.
44. Kapczinski F, Dias VV, Kauer-Sant'Anna M, et al. Clinical implications of a staging model for bipolar disorders. *Expert Rev Neurother*. 2009;9(7):957–966.
45. Naismith SL, Hermens DF, Ip TK, et al. Circadian profiles in young people during the early stages of affective disorder. *Transl Psychiatry*. 2012;2(5):e123.
46. Lagopoulos J, Hermens DF, Hatton SN, et al. Microstructural white matter changes are correlated with the stage of psychiatric illness. *Transl Psychiatry*. 2013;3(4):e248.
47. Penninx BWJH, 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.
48. Wiersma JE, Hovens JGFM, van Oppen P, et al. The importance of childhood trauma and childhood life events for chronicity of depression in adults. *J Clin Psychiatry*. 2009;70(7):983–989.
49. Hermens DF, Naismith SL, Lagopoulos J, et al. Neuropsychological profile according to the clinical stage of young persons presenting for mental health care. *BMC Psychol*. 2013;1(1):8.
50. Roca M, Gili M, Garcia-Garcia M, et al. Prevalence and comorbidity of common mental disorders in primary care. *J Affect Disord*. 2009;119(1–3):52–58.
51. Hopkinson G. The prodromal phase of the depressive psychosis. *Psychiatr Neurol (Basel)*. 1965;149(1):1–6.