

ORIGINAL ARTICLE

## A comparison of the Major Depression Inventory (MDI) and the Beck Depression Inventory (BDI) in severely depressed patients

ANASTASIOS KONSTANTINIDIS<sup>1</sup>, KLAUS MARTINY<sup>2</sup>, PER BECH<sup>2</sup> & SIEGFRIED KASPER<sup>1</sup>

<sup>1</sup>Department of Psychiatry and Psychotherapy, Division of Biological Psychiatry, Medical University Vienna, Austria, and

<sup>2</sup>Psychiatric Research Unit, Frederiksborg General Hospital, Frederiksborg, Denmark

### Abstract

**Background.** We set out to examine the psychometric properties of the MDI in comparison to the BDI in a mixed group of patients with primary depression. **Methods.** At the Department of Biological Psychiatry in Vienna currently depressed inpatients with either a depressive or a schizo-affective disorder filled out both MDI and BDI on day of admission and at a time-point two weeks later during their treatment. Furthermore the Hamilton Depression Scale (HAM-D) was administered by the treating clinician at both time-points. **Results.** In total, 51 patients were included in the study. The non-parametric item response analysis was preferred to the classical Cronbach coefficient  $\alpha$  as the latter is influenced by the number of items in a questionnaire. MDI obtained a Mokken analysis coefficient above 0.40, indicating unidimensionality. To determine external validity severely depressed patients with psychotic symptoms ( $N = 10$ ) were compared to the remaining non-psychotic depressed patients ( $N = 41$ ). Although BDI and MDI showed a lower score for psychotic than for non-psychotic inpatients, the standard deviations for both were greater for psychotic inpatients. On the intercorrelations between the different scales, MDI showed for all coefficients values above 0.70. On the other hand BDI and MDI both showed the same degree of linear relationship as the usual versions of HAM-D. **Conclusion.** Our results demonstrate that the MDI had the highest coefficients values and was sufficient as a measure for depressive disorders in psychiatric patients.

**Key Words:** Major Depression Inventory, Beck Depression Inventory, depression, validation

### Introduction

Over the last decades the use of patient-administered questionnaires for depression screening has increased, both in general population studies and in the general practice setting, to measure the prevalence of depression according to the International Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) or the Diagnostic and Statistical Manual of mental Disorders, 4th Edition (DSM-IV) [1–5].

The clinician-administered PRIME-MD [6] (a rating scale developed to screen for DSM-IV major depression), was subsequently adapted for self-administration by patients (Patient Health Questionnaire (PHQ-9)) [7]. The PHQ-9 was shown to have both an acceptable sensitivity and specificity for the recognition of DSM-IV major depression [8]. A version of the PRIME-MD in accordance with ICD-10 has been suggested but is still under development.

However, the Major Depression Inventory (MDI) was developed by a Danish WHO Collaborating Centre as a patient-administered questionnaire covering clinical validity of both ICD-10 depression and DSM-IV depression [9]. Sensitivity and specificity as well as responsiveness in antidepressant treatment have been found to be as high for the MDI as for the PHQ-9 [8–10]. In contrast to the Beck Depression Inventory (BDI) [11] the MDI has been used to analyse the prevalence of depression in the general population [5]. In this respect both the MDI and the PHQ-9 are superior to the BDI, because the clinical validity of the BDI when covering the ICD-10 or DSM-IV symptoms of depression is limited, but also because the BDI covers items which have been found problematic when contacting the general population [12].

The use of patient-administered questionnaires is, however, limited in severely depressed patients

Correspondence: A. Konstantinidis, Specialist in Psychiatry, Department of Psychiatry and Psychotherapy, Division of Biological Psychiatry, Medical University Vienna, Währinger Gürtel 18–20, A-1090 Wien, Austria. Tel: +43 1 40 400 3568. Fax: +43 1 40 400 3099. E-mail: anastasios.konstantinidis@meduniwien.ac.at

(Received 7 March 2010; accepted 7 July 2010)

due to the degree of dissimulation (denial of symptoms) and due to behavioural aspects of depressive states such as motor retardation. The Beck Depression Inventory was originally administered by an interviewer who was supposed to read each item aloud and ask the patient to select the statement that seemed most appropriately to capture his or her feelings [11]. In a previous Bech et al. study using the BDI, more than 95% of the scale items were also completed by inpatients [13].

On this background we wanted to retest the hypothesis that MDI is an adequate tool for detecting and evaluating the degree of depression in comparison to the BDI. For this purpose we evaluated the data of a mixed group of patients with primary depression, also including severely depressed inpatients with psychotic symptoms, who had filled out the MDI and BDI at least twice during their treatment in a time frame of 2 weeks, and who had been rated by a clinician using the Hamilton Depression Scale (HAM-D) [14,15]. Three of the various versions of the HAM-D were used: the conventional HAM-D<sub>17</sub> and HAM-D<sub>21</sub> as well as the Structured Interview Guide for the Hamilton Depression Scale-Seasonal Affective Disorders version (SIGH-SAD), which also includes psychotic depression items, as well as the HAM-D<sub>6</sub>, a six-item subscale of the HAM-D, shown in validation studies to be the psychometrically most valid measure of the severity of depressive states [16–18].

## Method

### *Study design and subjects*

MDI, BDI and the HAM-D (HAM-D<sub>17</sub>, HAM-D<sub>21</sub>, SIGH-SAD) are used on a regular basis in diagnosing and documenting the clinical benefit of ongoing treatment on depressive inpatients in the Division of Biological Psychiatry in the Medical University of Vienna.

We collected and evaluated the data of depressed inpatients with or without psychotic features and currently depressed inpatients with a schizo-affective disorder, who had filled out both MDI and BDI on the day of admission and at a time point 2 weeks afterwards, during their treatment in an inpatient ward at the Division of Biological Psychiatry in Vienna. Furthermore the interview for the HAM-D should have been administered by the treating clinician on both time points.

Depressive inpatients obviously unable to complete the questionnaires, e.g., not able to understand German, or patients with aphasia, and patients who had only filled out the ratings at one time point or at none of the time points were excluded.

The diagnosis of depression and the Hamilton Depression Scale interview were made without knowing the scores achieved on the self-administered questionnaires (MDI, BDI).

Treatment was administered by the treating clinicians according to clinical judgment and in accordance with published Consensus guidelines regarding the pharmacological treatment of depressive patients [19,20]. Treatment included the administration of Selective Serotonin Reuptake Inhibitors (SSRI), other classes of antidepressants, neuroleptics for psychotic symptoms, and the use of benzodiazepines as sedative/hypnotic treatment.

### *Measures*

*Self-assessment scales.* The Major Depression Inventory (MDI) [9] was used as a self-assessment scale. The MDI consists of 12 items scored on a frequency response scale from “none of the time” (zero) to “all of the time” (five). The score on this scale is calculated as the sum of the single items, but for the items *Increased/Reduced Appetite* and the items *Agitation/Retardation* only the highest score is used. Thus, the total score consists of 10 items and the score range is therefore from 0 to 50. The Beck Depression Inventory (BDI) [11] was used as a comparator. This scale consists of 21 items scored from 0 to 3 based on short items descriptors. The score on this scale is calculated as the sum of the single items and the score range is thus from 0 to 63 [11].

*Clinician reported scales.* The Hamilton Depression Rating Scale (HAM-D<sub>21</sub>) was used in its 21-item version [21] which includes the 17 items endorsed by Hamilton [22], supplemented by three items measuring psychotic dimensions of depression.

The Structured Interview Guide for the Hamilton Depression Scale-Seasonal Affective Disorders version (SIGH-SAD) [18] was used as a measure of atypical symptoms of depression, using the following items: “social withdrawal”, “increased appetite”, “carbohydrate craving”, “increased food intake”, “weight gain”, “increased sleep” and “fatigability”. These items are scored from 0 to 4 and the scale thus has a score range from 0 to 28.

### *Statistical analysis*

*Validation (internal validity).* The non-parametric item response analysis [23–25] was preferred to the classical Cronbach coefficient  $\alpha$  as the latter is influenced by the number of items in a questionnaire: The higher the number of items, the higher the  $\alpha$  coefficient will be, which is problematic when comparing scales with

such a large range of items (from 6 to 21). Moreover, Cronbach's coefficient  $\alpha$  is essentially a measure of reliability, i.e. the agreement between items in an attempt to make alternative versions of the individual items. As discussed by Greenberg [26] the Beck Depression Inventory contains many alternative items within the cognitive theory of depression, implying a psychometrical bias which the coefficient  $\alpha$  is not able to test. This was the problem with the Cronbach coefficient  $\alpha$  which was addressed by Loevinger when she developed the coefficient of homogeneity, which tests to what extent each item in a scale contributes with additional information about the dimension being tested so that the total score is a sufficient statistic [27]. Mokken [23] used the Loevinger coefficient of homogeneity in his program. The coefficient of homogeneity is independent of the number of items in the scale. According to Mokken a coefficient of homogeneity of 0.30–0.39 is only just acceptable, while a coefficient of 0.40 or higher is acceptable; indicating that the total score is a sufficient statistic.

*External validity.* When comparing differences between groups of patients, the non-parametric Mann–Whitney analysis was used [28] while the non-parametric Spearman test was employed when correlating the scales [28]. For standard deviation comparison, the *F*-test was used. The statistical level of significance was  $P \leq 0.05$ , two-tailed.

## Results

In total, 51 patients were included in the study (37 females and 14 males). The mean age of the patients was 54.4 years (14.8), with a range from 22 to 88 years of age. Using the ICD-10 [3], 10 patients were classified as suffering from recurrent depression episodes, current severe depression, and with psychotic mood-congruent syndromes. In the following, these 10 patients have been compared with the remaining 41 patients (three of whom were bipolar (two) or schizo-affective (one); 12 had a single depressive episode, and 26 had a recurrent unipolar depression without psychotic symptoms).

### Internal validity

Table I shows the coefficient of homogeneity for the various scales. The MDI obtained a coefficient above 0.40 for both psychotic and non-psychotic patients, indicating higher unidimensionality. The HAM-D<sub>6</sub> was just around 0.40, while all the other scales were below, but the coefficients were higher in the psychotic patients.

Table I. Mokken analysis at endpoint coefficient of homogeneity for the various scales.

Scales	Mokken analysis at endpoint coefficient of homogeneity		
	All patients ( <i>N</i> = 51)	Non-psychotic ( <i>N</i> = 41)	Psychotic ( <i>N</i> = 10)
HAM-D <sub>21</sub>	0.26	0.30	0.17
HAM-D <sub>17</sub>	0.28	0.31	0.22
HAM-D <sub>6</sub>	0.37	0.32	0.46
SIGH <sub>7</sub>	0.33	0.32	0.36
BDI <sub>21</sub>	0.27	0.25	0.38
MDI	0.60	0.58	0.72

### External validity

Table II shows the mean (SD) scores of the various scales. The severely depressed patients with psychotic symptoms (*N* = 10) are compared to the remaining depressed patients without psychotic symptoms (*N* = 41). The HAM-D shows that the psychotic patients scored numerically higher than the non-psychotic patients, but the difference was only statistically significant on the HAM-D<sub>21</sub>. Especially the sum of the extra items on HAM-D<sub>21</sub> from numbers 18 to 21 containing psychotic symptoms was significantly higher in the psychotic than in the non-psychotic patients ( $P \leq 0.01$ ). No difference between the two groups of patients was seen as to the sum of the HAM-D sleep items.

The self-administered questionnaires (both BDI and MDI) showed that the psychotic patients scored numerically lower than the non-psychotic patients, but this difference was not statistically significant (Mann–Whitney  $P \leq 0.05$ ). However, the standard deviations for both the BDI and the MDI were

Table II. The mean (SD) scores of the various scales.

Scales	ICD-10 depression	
	Moderate/ severe without psychotic symptoms ( <i>N</i> = 41)	Severe with psychotic symptoms ( <i>N</i> = 10)
HAM-D <sub>21</sub>	23.8 (4.7)	26.3 (3.7)*
HAM-D <sub>17</sub>	22.1 (4.1)	22.3 (3.8)
HAM-D <sub>6</sub>	11.9 (1.6)	12.5 (1.8)
HAM-D <sub>18–21</sub>	1.7 (1.1)	4.0 (1.6)**
HAM-D <sub>suicide3</sub>	0.9 (0.6)	1.2 (0.4)
HAM-D <sub>sleep4–6</sub>	3.5 (1.9)	3.1 (2.0)
SIGH <sub>7</sub>	6.4 (2.0)	6.5 (2.8)
BDI <sub>21</sub>	28.2 (6.5)	26.5 (9.0)
MDI	34.5 (6.1)	30.3 (10.4)†

\* $P \leq 0.05$  for mean score difference (Mann–Whitney);

\*\* $P \leq 0.01$ .

† $P \leq 0.05$  for standard deviation (F-test).

numerically greater for the psychotic patients compared to the non-psychotic patients, a tendency not seen for the Hamilton Depression Scales. For the MDI, the *F*-test showed that this difference in standard deviation was statistically significant (Table II).

Table III shows the percentage of improvement after two weeks of therapy. Non-psychotic patients had an improvement percentage of above 25% on the HAM-D as well as the self-rating scales, but the differences between the scales were not statistically significant. However, the HAM-D<sub>18-21</sub> item improvement (57.5%) in the psychotic patients was significantly greater than the improvement found in non-psychotic patients (26.1%) ( $P \leq 0.01$ ). The suicide item as well as the sleep items obtained the highest improvement percentage in both groups of patients over the 2 weeks of therapy compared to the other scales.

Table IV shows the intercorrelations between the different scales. For the MDI, all coefficients were above 0.70. For the other scales the coefficients were slightly below 0.70 when correlated with SIGH<sub>7</sub>. Furthermore BDI and MDI showed the same degree of linear relationship with the usual versions of the HAM-D.

## Discussion

In one of the first trials in outpatients with neurotic depression in which both the HAM-D and BDI was used to measure outcome of imipramine versus cognitive psychotherapy, Beck and his group [29] obtained a baseline score of approximately 21 on the HAM-D and 31 on the BDI. However, in outpatients with major depression the baseline scores before treatment on these scales are typically an HAM-D score around 24 and a BDI score around 28 [30].

Table III. Percentage of improvement after two weeks of therapy as measured by the various scales.

Scales	Percentage improvement in 2 weeks therapy	
	Mild/ moderate depression ( <i>N</i> = 41)	Severe (psychotic) ( <i>N</i> = 10)
HAM-D <sub>21</sub>	37.8%	38.4%
HAM-D <sub>17</sub>	38.5%	34.9%
HAM-D <sub>6</sub>	28.6%	23.2%
HAM-D <sub>18-21</sub>	26.1%	57.5%**
HAM-D <sub>suicide3</sub>	55.6%	66.7%
HAM-D <sub>sleep4-6</sub>	82.9%	83.9%
SIGH <sub>7</sub>	38.8%	23.1%
BDI <sub>21</sub>	37.2%	32.1%
MDI	35.7%	31.4%

\*\* $P \leq 0.01$ .

Table IV. Intercorrelation between the scales; Spearman coefficients (all patients *N* = 51) endpoint.

	HAM-D <sub>21</sub>	HAM-D <sub>17</sub>	HAM-D <sub>6</sub>	SIGH <sub>7</sub>	BDI <sub>21</sub>	MDI
HAM-D <sub>21</sub>	1.00					
HAM-D <sub>17</sub>	0.98	1.00				
HAM-D <sub>6</sub>	0.91	0.90	1.00			
SIGH <sub>7</sub>	0.68	0.68	0.63	1.00		
BDI <sub>21</sub>	0.86	0.86	0.77	0.69	1.00	
MDI	0.82	0.82	0.75	0.72	0.82	1.00

Among the group of patients in our study with moderate to severe depression, but without psychotic symptoms according to ICD-10, the baseline score on HAM-D and BDI was within the expected scores. In our group of depressed patients with psychotic symptoms the BDI score was even below 28 at baseline, indicating the relatively high BDI scores obtained in patients for cognitive therapy reflect that the BDI includes more cognitive-related items. The scores on the MDI and HAM-D<sub>17</sub> and their relative magnitude in our study are in agreement with other studies using these two scales<sup>31</sup>.

As in agreement with Sayers et al. [32], in our group of patients with psychotic symptoms we found out that the standard deviation of the BDI or MDI was higher than for the HAM-D when compared to the non-psychotic depressive group of patients. The psychometric analysis of both the BDI and the MDI showed that even in the psychotic group of patients the total score was a sufficient statistic. However, the coefficient of homogeneity for BDI was only just acceptable (0.38) while it was quite acceptable (0.72) for MDI.

The specific core items of depression as captured by the HAM-D<sub>6</sub> obtained the highest coefficient of homogeneity, indicating the importance of these items in diagnosing depression. When evaluating the various scales' sensitivity to measure improvement during the first 2 weeks of therapy, the HAM-D items containing the psychotic symptoms were the ones to discriminate significantly between the psychotic and the non-psychotic group of patients. In general, however, it was the HAM-D subscale of insomnia that indicated the greatest improvement (approximately 83%), but also the item "suicidal thoughts" showed a relatively high improvement (approximately 60%). This effect can be explained by co-medication with benzodiazepines in the first 2 weeks of treatment with antidepressants (mainly SSRIs).

The HAM-D<sub>17</sub>, the BDI and the MDI had a high correlation (36%) in showing improvement over the first 2 weeks of therapy. This result demonstrates the ability of self-rating scores to detect achieved improvements during therapy. Furthermore, an improvement

of 25% was seen on the specific depression symptoms, a percentage usually to be found in studies with antidepressants [33].

The MDI presented the highest coefficient of homogeneity out of all the scales used in our trial using the non-parametric item response analysis and documented improvement due to adjusted psycho-pharmacotherapy equally as well as an observer rating tool (HAM-D). The MDI provides a new, easy to use self rating scale, with fewer items, rendering it faster to complete. Although it was not constructed to diagnose psychotic features in depressive patients, patients belonging to this group could be diagnosed as depressed using the MDI. The development of an MDI adapted to take into account psychotic features is now under consideration. Further studies could evaluate the use of MDI as a measure in the assessment of remission in patients with a major depressive disorder undergoing drug therapy.

Short and valid measurement tools may lead to better diagnosis and treatment of depression thus reducing the suffering of the patients and the associated costs of treatment.

### Keypoints

- Major Depression Inventory (MDI) presented the highest coefficient of homogeneity out of all the scales used in our trial (BDI – Beck Depression Inventory; HAM-D – Hamilton Depression Rating Scale) and documented improvement due to adjusted psycho-pharmacotherapy equally as well as an observer rating tool (HAM-D)
- The psychometric analysis of both the BDI and the MDI showed that even in the psychotic group of patients the total co-efficiency score was sufficient. However, the coefficient of homogeneity for BDI was only just acceptable (0.38) while it was quite acceptable (0.72) for MDI
- Our study suggests that the MDI could be a useful tool in diagnosing and monitoring depressive patients

### Acknowledgements

None.

### Statement of interest

Dr Konstantinidis has received honoraria from Pfizer and AstraZeneca, served as consultant for AstraZeneca, and as a speaker for AstraZeneca and Bristol Myers Squibb. Dr Martiny declares no conflicts of interests. Dr Bech has occasionally received funding

from and been speaker or member of advisory boards for pharmaceutical companies with an interest in drug treatment of affective disorders (Astra-Zeneca, Lilly, H. Lundbeck A/S, Lundbeck Foundation, Organon). Dr Kasper has received grant/research support from Eli Lilly, Lundbeck, Bristol-Myers Squibb, GlaxoSmithKline, Organon, Sepracor and Servier; has served as a consultant or on advisory boards for AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, Eli Lilly, Lundbeck, Pfizer, Organon, Schwabe, Sepracor, Servier, Janssen, and Novartis; and has served on speakers' bureaus for AstraZeneca, Eli Lilly, Lundbeck, Schwabe, Sepracor, Servier, Pierre Fabre, and Janssen.

### References

- [1] APA. Diagnostic and statistical manual of mental disorders. 4th ed. Text revision. Washington, DC : American Psychiatric Association; 2000.
- [2] Fink P, Jensen J, Borgquist L, Brevik JI, Dalgard OS, Sandager I, Engberg M, et al. Psychiatric morbidity in primary public health care: a Nordic multicentre investigation. Part I: method and prevalence of psychiatric morbidity. *Acta Psychiatr Scand* 1995;9:409–18.
- [3] WHO. International classification of diseases and related health problems. 10th revision. Geneva: World Health Organization; 1993.
- [4] Krause P, Wittchen HU, Höfler M, Winter S, Spiegel B, Pfister H. Generalized anxiety and depression in primary care. Study design and methods. *Fortschr Med Orig* 2001; 119(Suppl 1):5–12.
- [5] Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychol Med* 2003;33:351–6.
- [6] Spitzer RL, Williams JBW, Kroenke K, Linzer M, deGruy FV 3rd, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *J Am Med Assoc* 1994;272:1749–56.
- [7] Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD. The PHQ primary care study. *J Am Med Assoc* 1999;282:1737–44.
- [8] Löwe B, Schenkel I, Carney-Doebbeling C, Gobel C. Responsiveness of the PHQ-9 to Psychopharmacological Depression Treatment. *Psychosomatics* 2006;47:62–7.
- [9] Bech P, Rasmussen NA, Olsen LR, Noerholm V, Abildgaard W. The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. *J Affect Disord* 2001;66:159–64.
- [10] Bech P, Lunde M, Undén M. Social Adaptation Self-evaluation Scale (SASS): Psychometric analysis as outcome measure in the treatment of patients with major depression in the remission phase. *Int J Psychiatry Clin Pract* 2002; 6:141–6.
- [11] Beck AT, Ward CH, Mendelson M. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561–71.
- [12] Naughton MJ, Wiklund I. A critical review of dimension-specific measures of health-rated quality of life in cross-cultural research. *Qual Life Res* 1993;2:397–432.
- [13] Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG. Quantitative rating of depressive states. *Acta Psychiatr Scand* 1975;51:161–70.

- [14] Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.
- [15] Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967;6:278–96.
- [16] Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: Has the gold standard become a lead weight? *Am J Psychiatry* 2004;161:2163–77.
- [17] Licht RW, Qvitzau S, Allerup P, Bech P. Validation of the Bech-Rafaelsen Melancholia Scale and the Hamilton Depression Scale in patients with major depression: is the total score a valid measure of illness severity? *Acta Psychiatr Scand* 2005;111:144–9.
- [18] Williams JBW, Link MJ, Rosenthal NE, Amira L, Terman M. Structured Interview Guide for the Hamilton Depression Rating Scale – Seasonal Affective Disorder Version (SIGH-SAD) (revised ed.) New York: New York State Psychiatric Institute; 1994.
- [19] Bauer M, Bschor T, Pfennig A, Whybrow PC, Angst J, Versiani M, Möller HJ on behalf of the WFSBP Task Force on Treatment Guidelines for Unipolar Depressive Disorders. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders in Primary Care. *World J Biol Psychiatr* 2007;8:67–104.
- [20] Kasper S, Lehofer M, editors. Depression – Medikamentöse Therapie. Konsensus-Statement – State of the Art 2007. *CliniCum Psy Sonderausgabe 01 2007*, Vienna.
- [21] Bech P. Rating scales for psychopathology, health status and quality of life. A compendium on documentation in accordance with the DSM-III-R and WHO systems. Berlin: Springer; 1993.
- [22] Williams JBW. A structured interview guide for the Hamilton depression rating scale. *Arch Gen Psychiatry* 1988;45:742–7.
- [23] Mokken RJ. A theory and procedure of scale analysis. Paris: Mouton; 1971.
- [24] Molenaar IW, Debets P, Systma K, Hemker BT. User's manual MSP: A program for the Mokken scale analysis for polytomous items (version 3.0). Groningen: lec ProGamma; 1994.
- [25] Sijtsma K, Molenaar IW Introduction to nonparametric item response theory. London: Sage; 2002.
- [26] Greenberg G. Manufacturing depression. The secret history of a modern disease. London: Bloomsbury; 2010.
- [27] Loevinger J. Objective tests as instruments of psychological theory. *Psychol Rep* 1957;3:635–94.
- [28] Siegel S. Nonparametric statistics. New York: McGraw-Hill; 1956.
- [29] Rush AJ, Beck AT, Kovacs M, Hollon S. Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cogn Ther Res* 1977;1: 17–37.
- [30] Avery DH, Holzheimer PE, Fawaz W, Russo J. A controlled study of repetitive transcranial magnetic stimulation in medication-resistant major depression. *Biol Psychiatry* 2006;59: 187–94.
- [31] Martiny K. Adjunctive bright light in non-seasonal major depression. *Acta Psychiatr Scand Suppl* 2004;425:7–28.
- [32] Sayer NA, Sackheim HA, Moeller JR, Prudic J, Devanand DP, Coleman EA, Kiersky JE, American Psychological Association. The relations between observer-rating and self-report of depressive symptomatology. *Psychol Assess* 1993;5: 350–60.
- [33] Bech P. Pharmacological treatment of depressive disorders: a review. In: Maj M, Sartorius N, editors. Depressive disorders. WPA Series Evidence and Experience in Psychiatry. 2nd ed. Chichester: John Wiley & Sons ; 2002. p. 89–127.