

Arabic version of the Major Depression Inventory as a diagnostic tool: reliability and concurrent and discriminant validity

M.H. Fawzi,¹ M.M. Fawzi¹ and W. Abu-Hindi¹

الإخراجة العربية من جرد المعطيات حول الاكتئاب الجسيم باعتبارها أداة تشخيصية: المَعَوَّلِيَّة والصَّوَابِيَّة التَّمْيِيزِيَّة والتَّزَامِيَّة
منير فوزي، مهاب فوزي، وائل أبو هندي

الخلاصة: يُعدُّ جرد المعطيات حول الاكتئاب الجسيم MDI بمثابة استبيان موجز يستخدم لتقييم وجود اضطراب اكتئابي. وقد أعدَّ الباحثون إخراجة عربية من جرد المعطيات حول الاكتئاب الجسيم وقاموا باختبار مَعَوَّلِيَّتِهَا وصَوَابِيَّتِهَا التَّمْيِيزِيَّة والتَّزَامِيَّة باعتبارها أداة تشخيصية للاضطراب الاكتئابي الجسيم. وشملت الدراسة خمسين مريضاً من مرضى العيادات الخارجية في مصر ممن يعانون من الاضطراب الاكتئابي الجسيم (بعد تقييمهم سريريًا ومن خلال المقابلات السريرية البنوية لتجري اضطرابات المحور I DSM-IV-TR) ثم قارن الباحثون النتائج بما حصلوا عليه من خمسين شاهداً صحيحاً باستخدام جرد المعطيات حول الاكتئاب الجسيم، وجرد المعطيات حول الاكتئاب بحسب "بيك"، وجرد المعطيات حول القَلَق بحسب "سبيلبرغر". واتضح للباحثين أن قيمة معامل كرونباخ ألفا 0.91، وأن قيمة معامل الترابط بين الفئات 0.98 (بفاصلة ثقة 95%: 0.97 - 0.99). كما لوحظ ترابط قوي بين أحراز جرد المعطيات حول الاكتئاب الجسيم وأحراز جرد المعطيات حول الاكتئاب بحسب "بيك"، ($r=0.81$)، في حين لم يكن هناك ترابط يُعْتَدُّ به مع أحراز جرد المعطيات حول القَلَق بحسب "سبيلبرغر". كما أنه باستخدام خوارزمية جرد المعطيات حول الاكتئاب الجسيم كانت الحساسية 88.4% والنوعية 78.9%. وخلص الباحثون إلى أن الإخراجة العربية من جرد المعطيات حول الاكتئاب الجسيم تتمتع بمَعَوَّلِيَّة ممتازة وبدرجة مقبولة من الصَّوَابِيَّة التَّمْيِيزِيَّة والتَّزَامِيَّة.

ABSTRACT The Major Depression Inventory (MDI) is a brief questionnaire to assess the presence of a depressive disorder. We prepared an Arabic version of the MDI and tested its reliability and concurrent and discriminant validity as a diagnostic tool of major depressive disorder. A group of 50 Egyptian outpatients with major depressive disorder (assessed clinically and with the Structured Clinical Interview for DSM-IV-TR Axis I Disorders) were compared with 50 healthy controls using the MDI-A, Beck Depression Inventory (BDI) and Spielberger State-Trait Anxiety Inventory (STAI). Cronbach α was 0.91 and intraclass correlation coefficient was 0.98 (95% CI: 0.97-0.99). Scores on the MDI-A strongly correlated with BDI scores ($r=0.81$) but insignificantly correlated with STAI scores. Using the MDI scoring algorithm, the sensitivity was 88.4% and specificity 78.9%. We conclude that the MDI-A has an excellent reliability and an acceptable concurrent and discriminant validity.

Version en langue arabe du *Major Depression Inventory* en tant qu'outil diagnostic : fiabilité et validité concurrente et discriminante

RÉSUMÉ Le *Major Depression Inventory* est un questionnaire abrégé destiné à évaluer la présence d'un trouble dépressif. Nous avons élaboré une version en langue arabe de ce questionnaire et avons testé sa fiabilité et sa validité concurrente et discriminante en tant qu'outil diagnostic d'un trouble dépressif majeur. Un groupe de 50 patients égyptiens soignés en consultation externe pour un trouble dépressif majeur (évalués cliniquement et par une entrevue cliniquement structurée pour les troubles de l'Axe I du DSM-IV-TR) a été comparé à un groupe témoin de 50 personnes en bonne santé en utilisant la version en langue arabe du *Major Depression Inventory*, ainsi que l'Inventaire de la dépression de Beck et l'Inventaire d'anxiété état-trait de Spielberger. Le coefficient α de Cronbach était de 0,91 et le coefficient de corrélation intraclass était de 0,98 (IC à 95 % : 0,97-0,99). Les scores obtenus au *Major Depression Inventory* étaient fortement corrélés à ceux de l'Inventaire de la dépression de Beck ($r=0,81$) mais étaient faiblement corrélés aux résultats pour l'Inventaire d'anxiété état-trait de Spielberger. La sensibilité était de 88,4 % et la spécificité de 78,9 % selon l'algorithme de notation du *Major Depression Inventory*. Nous en concluons que la version arabe du *Major Depression Inventory* est hautement fiable et possède une validité concurrente et discriminante acceptable.

¹Department of Psychiatry, Faculty of Medicine, Zagazig University, Zagazig, Egypt (Correspondence to M.H. Fawzi: mounir.fawzi40@gmail.com).

Received: 08/09/10; accepted: 22/09/10

Introduction

One of the problems with the diagnosis of major depressive disorder, at least in certain countries, is connected with the lack of appropriate instruments for objective and systematic evaluation of depression [1]. Many of the existing instruments in common use to diagnose major depressive disorder belong to the era before development of the *Diagnostic and statistical manual of mental disorders*, 4th edition (DSM-IV), although some of them have been modified to have a closer coverage of the DSM-IV criteria for major depression [2].

The Major Depression Inventory (MDI), developed by Bech et al. for the World Health Organization in 2001 [3], was based on the DSM-IV symptoms of major depression and the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) category of moderate to severe depression which includes both duration and intensity criteria. The MDI can be used as a diagnostic instrument in which the items are dichotomized to indicate the presence or absence of each of the symptoms. This instrument is brief, allows clinicians to assess the presence of a depressive disorder and can be filled in by the patient him/herself [4]. In its original English version, the inventory is reported to have internal and external validity higher than for Zung Depression Rating Scale [5].

It cannot be overestimated that there is a need for equivalent language versions of instruments assessing depression to undertake multi-centre research and to obtain meaningful comparison of results obtained in different countries. The original English language version of the MDI has been translated into several languages, including Danish, Dutch, Finnish, French, German, Swedish, Turkish, Spanish and Serbian, but the reliability and validity of the MDI translated versions have been tested in only a limited number of studies,

and to date there has been no validated version available in Arabic language.

The objectives of this study were to translate and adapt the MDI into Arabic, to test the reliability of the resulting version and to evaluate the concurrent and discriminant validity of the MDI algorithm translated version on a group of Egyptian outpatients with major depression and a group of healthy controls.

Method

Subjects

Using *Epi-Info*, version 6 software (*Statcalc* module), the sample size was calculated to be 49 participants in each group, based on $\alpha = 5\%$ and $\beta = 10\%$ to detect correlation coefficients of 0.40 to 0.60.

We consecutively recruited 50 depressed patients, between October 2009 and March 2010, from all patients with a diagnosis of depression who were referred for consideration for study inclusion. Referrals were made by the attending staff at the psychiatric outpatient clinics of the Zagazig university hospitals and referrals from a private psychiatric clinic of an experienced psychiatrist in Zagazig, Egypt. For inclusion in the study, patients were required to be able to at least read and write, to be in the age range 18–60 years and to fulfil DSM-IV, Text Revision (DSM-IV-TR) criteria for major depressive disorder. We excluded patients with any current or previous DSM-IV-TR primary axis I diagnosis other than major depressive disorder and patients who were unable to be interviewed. Patients with a concomitant serious medical illness, including any cardiovascular, hepatic, renal, respiratory, haematological, endocrinological or neurological disease, were also ineligible.

A further 50 healthy, age- and sex-matched volunteers were enrolled as controls from the companions of patients ($n = 36$) and from service

personnel ($n = 14$). Freedom from past or current physical or psychiatric illness was confirmed by medical evaluation and results of the Structured Clinical Interview for DSM Disorders (SCID).

Instruments

Major Depression Inventory, Arabic version (MDI-A)

The MDI used here was the Arabic translated version (MDI-A). A copy of the instrument is available from the authors. The original MDI is a self-rating inventory using the past 2 weeks as the time frame. In principle, MDI is composed of 12 items, but because items 8 and 10 are composed of 2 sub-items each (a and b) with only the highest score of either a or b is countable, the MDI has, functionally, only 10 items. Responses are scored on a 6-point Likert-type scale, ranging from 0 (at no time) to 5 (all the time) with a potential range of 0–50. In addition to the calculation of the severity of depression, MDI can be scored in an algorithmic way to diagnose major depression. In this case, the items are dichotomized to indicate the presence or absence of each of the symptoms. The number of items is reduced to 9 by combining items 4 and 5 (which belong to the same category in DSM-IV) and considering only the item with the higher score. A score of 4 or 5 (that is, most of the time or all the time) is required on 5 of the 9 items, but at least 1 of these 5 items must be either item 1 (depressed mood) or item 2 (loss of interest) [3].

The translation procedure followed the recommendations by Sartorius and Kuyken [6]. Two psychiatrists whose mother tongue was Arabic, but also fluent in English, translated independently the English MDI version into modern standard Arabic. An Egyptian panel of 3 psychiatrists and 2 translators reviewed the 2 translations to assess cultural and linguistic appropriateness, resolve discrepancies and merge the 2 translations into 1 reconciled version. The resulting version was then translated back into

English by 2 other English-speaking language specialists who had no medical or clinical background and had not seen the original MDI before. We compared the back-translated versions with the original, and discrepancies were addressed to produce the final version which was tested for understandability and acceptability in a small pilot study involving a convenience sample of 5 patients with major depression and 5 relatives with different levels of education, none of whom was illiterate. These participants were not included in the main study but also gave informed consent.

Other instruments

- Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version (SCID-I) [7]. This was used to exclude any current or lifetime psychiatric disorder other than major depression in the patient group. The mood disorders module of SCID was selected to serve as the clinical gold standard for assessing depression and was administered by trained psychiatrists.
- Beck Depression Inventory (BDI). This is a well-known self-report instrument. Its original version (21 items) was introduced in 1961 [8] and its reliability and validity have been established across a broad spectrum of clinical and non-clinical populations [9]. We used the most commonly used cut-off scores for BDI of < 10 to identify people within the normal range and > 16 to indicate clinical depression [10]. The Arabic version of the BDI by Abdel-Khalek [11] was used in this study.
- Spielberger's State-Trait Anxiety Inventory (STAI) [12]. The Arabic version of the STAI by Abdel-Khalek [13] was also used in this study.

Data collection

Ethical approval for the study was obtained from the research ethics committee at the Faculty of Medicine, Zagazig University, and written

informed consent was obtained from every participant.

All interviews were conjointly conducted by one of the authors and one of the 22 psychiatric staff colleagues who were trained in the use of the study instruments. They were found to have acceptable inter-rater reliabilities. All self-report inventories were filled out on the same day following the interview. To assess test–retest reliability, the MDI-A inventory was re-administered 2 weeks after the initial administration to 47 patients and 45 control subjects. Patients included in the re-test were only those who reported “no change since initial assessment”. To minimize the practice effect, the order of the items was reshuffled for the repeat test.

Statistical analysis

Demographic data were presented using descriptive statistics [mean and standard deviation (SD) for quantitative variables and frequencies for qualitative variables]. Differences between patients and controls were examined using the chi-squared test for categorical and *t*-test for continuous variables. Spearman rho (ρ) was used to assess correlations of background characteristics with MDI-A scores. MDI-A item scores of patients and controls were compared using the Mann–Whitney U non-parametric test.

The internal consistency reliability of the MDI-A scale was evaluated based on Cronbach for the whole scale, as well as for all but 1 items if any one was deleted (“if item deleted”). The correlation of each item score to the total scale score (“corrected item–total correlation”) was also measured through Pearson correlation coefficients.

Repeatability (test–retest reliability) was assessed by examining the rank order stability of MDI-A scores between the 2 testings. Using the procedure described by Shrout and Fleiss [14] a 2-way random effects model (absolute agreement) average measures intra-class correlation coefficient (ICC)

was used to calculate the test–retest reliability of the MDI-A. Based on the guidelines by Fleiss [15] an ICC value < 0.4 was considered to represent poor reliability, values 0.4–0.75 fair to good reliability and 0.75–1.0 excellent reliability. An ICC value of 1.0 represented perfect reliability. Test–retest reliability was, in addition, computed by Pearson product moment correlations. Concurrent and discriminant validity of the MDI-A were then evaluated using again the Pearson product moment correlations.

All statistical analyses were performed using SPSS, version 11.5 for Windows, and significance for all analyses was taken at the 5% level.

Results

The mean age of the patients with major depression was 39.5 years and 62% were female. Around half of patients were from low or very low social class and about more than one-third were not educated beyond basic level (Table 1). All participants were from Sharkia governorate. There were no significant differences between the background characteristics of patients and controls.

Using Spearman ρ , a positive correlation on the border of statistical significance was found between age and total MDI-A scores. Correlations between other background characteristics and total MDI-A scores were not significant (Table 1).

The mean total MDI-A score at baseline assessment was significantly higher in the patient group [33.2 (SD 4.3)] than the control group [9.5 (SD 8.7)] ($P < 0.001$). The mean of each individual MDI-A item at baseline are shown in Table 2 for both groups. All items, except item 10b, exhibited higher scores in the patient group than the control group (Table 2). In the patient group, the MDI-A items with the highest mean score were item 1 “Have you felt low in spirits or sad?” [4.1 (SD 0.9)] and item 2 “Have you lost interest in

Table 1 Background characteristics of patients with major depression and non-depressed controls and their correlation with total scores on the Depression Inventory, Arabic version (MDI-A)

Variable	Patients (n = 50)	Controls (n = 50)	χ^2 -test	Total (n = 100)	Correlation with MDI-A total scores	
	No.	No.	P-value	No.	Spearman ρ	P-value
Sex						
Male	19	20		39	0.176	0.08
Female	31	30	0.84	61		
Age (years)						
≥ 40	32	32	1.00	64	0.197	0.049
< 40	18	18		36		
Mean (SD)	39.6 (7.1)	40.9 (8.5)	0.41 ^a	42.0 (8.6)		
Social class						
High/middle	24	22		46	0.071	0.48
Low/very low	26	28	0.69	54		
Marital status						
Unmarried	21	24		45	0.031	0.76
Married	29	26	0.55	55		
Education						
Basic	18	19		57	0.649	0.42
Higher than basic	32	31	0.84	43		
Work status						
Employed	27	29		56	0.169	0.09
Unemployed	23	21	0.69	44		
Residence						
Rural	18	16		34	0.040	0.69
Urban	32	34	0.67	66		

^at-test.

SD = standard deviation.

your daily activities?" [4.1 (SD 0.8)]. The lowest scoring item was item 10b "Have you suffered from increased appetite?" [0.8 (SD 1.6)].

Reliability and repeatability were assessed. The MDI-A scale total score had an excellent reliability as indicated by Cronbach coefficient $\alpha = 0.91$, ICC = 0.98 (95% CI: 0.97–0.99) and the Spearman rank correlation coefficient $\rho = 0.95$ ($P < 0.001$). As shown in Table 3, apart from the corrected item–total correlation for item 10b: "Have you suffered from increased appetite?", which was low (–0.02), corrected item–total correlations ranged from 0.37 (item 8b: "Have you felt subdued or slowed down?") to 0.85 (item 2: "Have you lost interest in your daily activities?"). Cronbach alphas when a

particular item was deleted ranged from 0.89 to 0.92.

In this study we also assessed the concurrent and discriminant validity of the MDI-A. Results are shown in Table 4. Scores on the MDI-A strongly correlated with the BDI scores ($r = 0.81$, $P < 0.001$). In contrast, correlations of scores on the MDI-A with scores on the STAI-S and STAI-T were not significantly correlated [$r = 0.17$, $P = 0.08$ and $r = 0.19$, $P = 0.06$ respectively].

Using the cutoff scoring algorithm of the MDI, it was possible to correctly identify 38 out of 50 patients as having major depression and 45 out of 50 controls as not having major depression (Table 5). The diagnostic testing measures derived from this 2 × 2 contingency

table are shown on Table 6. Notably, the MDI-A had a sensitivity of 88.4% and specificity of 78.9%.

Discussion

We translated the MDI into a simple and acceptable modern standard Arabic language, which we preferred over a colloquial language, because it would be more likely to be understandable to people from all Arab countries. However, when translation produced words which were considered "difficult", i.e. not currently in common use in Egypt, we changed these words for more colloquial words or phrases, with the objective of not getting just word for-word equivalence but a cultural equivalence of the original English version. The fact

Table 2 Comparison of the of Major Depression Inventory, Arabic version (MDI-A) mean item scores between patients and controls at baseline assessment for patients with major depression and non-depressed controls

Item #	MDI-A item	Patients (n = 50)	Controls (n = 50)	Mann-Whitney test	
		Mean (SD)	Mean (SD)	z	P-value
1	Have you felt low in spirits or sad?	4.1 (0.9)	1.8 (1.6)	6.611	< 0.001
2	Have you lost interest in your daily activities?	4.1 (0.8)	1.5 (1.5)	7.404	< 0.001
3	Have you felt lacking in energy and strength?	4.0 (0.8)	1.5 (1.5)	7.255	< 0.001
4	Have you felt less self-confident?	3.7 (0.8)	1.1 (1.2)	7.655	< 0.001
5	Have you had a bad conscience or feelings of guilt?	2.8 (1.6)	0.5 (1.2)	6.675	< 0.001
6	Have you felt that life wasn't worth living?	2.1 (1.3)	0.4 (0.9)	6.197	< 0.001
7	Have you had difficulty in concentrating, e.g. when reading the newspaper or watching television?	3.0 (1.3)	0.6 (1.0)	7.275	< 0.001
8a	Have you felt very restless?	1.9 (2.0)	0.3 (0.6)	4.116	< 0.001
8b	Have you felt subdued or slowed down?	1.6 (1.8)	0.3 (0.8)	3.850	< 0.001
9	Have you had trouble sleeping at night?	2.8 (1.5)	0.5 (0.9)	6.852	< 0.001
10a	Have you suffered from reduced appetite?	2.5 (1.7)	0.5 (0.9)	5.576	< 0.001
10b	Have you suffered from increased appetite?	0.8 (1.6)	0.4 (1.1)	1.130	0.260
	Total score	33.2 (4.3)	9.5 (8.7)	8.480	< 0.001

SD = standard deviation.

Table 3 Reliability coefficients for items of the Major Depression Inventory, Arabic version (MDI-A)

Item #	MDI-A item	Internal consistency reliability		Test-retest reliability		
		Cronbach α		ICC (95% CI)	Spearman ρ	P-value
		Corrected item-total correlation	If item deleted			
1	Have you felt low in spirits or sad?	0.82	0.89	0.93 (0.89–0.95)	0.81	< 0.001
2	Have you lost interest in your daily activities?	0.85	0.89	0.94 (0.76–0.95)	0.81	< 0.001
3	Have you felt lacking in energy and strength?	0.84	0.89	0.93 (0.89–0.96)	0.84	< 0.001
4	Have you felt less self-confident?	0.81	0.89	0.96 (0.93–0.97)	0.91	< 0.001
5	Have you had a bad conscience or feelings of guilt?	0.71	0.90	0.97 (0.96–0.98)	0.94	< 0.001
6	Have you felt that life wasn't worth living?	0.67	0.90	0.98 (0.97–0.99)	0.96	< 0.001
7	Have you had difficulty in concentrating, e.g. when reading the newspaper or watching television?	0.79	0.89	0.97 (0.95–0.98)	0.94	< 0.001
8a	Have you felt very restless?	0.41	0.91	0.99 (0.98–0.99)	0.95	< 0.001
8b	Have you felt subdued or slowed down?	0.37	0.91	0.99 (0.98–0.99)	0.96	< 0.001
9	Have you had trouble sleeping at night?	0.70	0.90	0.99 (0.98–0.99)	0.98	< 0.001
10a	Have you suffered from reduced appetite?	0.62	0.90	0.96 (0.94–0.98)	0.92	< 0.001
10b	Have you suffered from increased appetite?	-0.02	0.92	0.97 (0.95–0.98)	0.94	< 0.001
	Total score	0.83	0.89	0.98 (0.97–0.99)	0.95	< 0.001

ICC = intraclass correlation coefficient; CI = confidence interval.

Table 4 Correlation matrix for the Major Depression Inventory, Arabic version (MDI-A), the Beck Depression Inventory (BDI), the State-Trait Anxiety Inventory, State (STAI-S) and the State-Trait Anxiety Inventory, Trait (STAI-T)

Instrument	MDI-A	BDI	STAI-S	STAI-T
MDI-A	1.00	-	-	-
BDI	0.81 ^a	1.00	-	-
STAI-S	0.17	0.19	1.00	-
STAI-T	0.19	0.20	0.25 ^b	1.00

^aCorrelation significant at the 0.01 level (2-tailed); ^bCorrelation significant at the 0.05 level (2-tailed).

Table 5 Comparison of major depression groupings generated by the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, research version (SCID-I) and the Major Depression Inventory, Arabic version (MDI-A) for patients with major depression and non-depressed controls

MDI-A classification	SCID classification		
	Patients	Controls	Total
	No.	No.	No.
Patients	38	12	50
Controls	5	45	50
Total	43	57	100

that all participants answered all the questions of the MDI-A indicates that this version is easy to understand and culturally appropriate. Participants were selected from both public and private psychiatric services so that people from diverse demographic and socioeconomic backgrounds were represented.

A limited number of studies previously tested the reliability and validity of

the MDI [15]. These studies, however, generally reported the reliability of MDI as very high, with Cronbach α ranging from 0.94 [4] to 0.89 [16]. Our finding of a Cronbach α of 0.91 is consistent with these studies and indicative of excellent internal consistency reliability of the items of MDI-Arabic translated version. Our finding of a test-retest reliability coefficient of 0.95, which is also

Table 6 Diagnostic testing measures for the Major Depression Inventory, Arabic version (MDI-A) derived from Table 5

Measure	Value (95% CI)
Overall accuracy (overall fraction correct) (%)	83.0 (0.75-0.88)
Misclassification rate (1-overall fraction correct) (%)	17.0 (0.12-0.25)
Sensitivity (%)	88.4 (0.79-0.94)
Specificity (%)	78.9 (0.72-0.84)
Positive predictive value	0.76 (0.68-0.81)
Negative predictive value	0.90 (0.82-0.95)
Positive likelihood ratio [sensitivity/(1-specificity)]	4.20 (2.78-5.73)
Negative likelihood ratio [(1-sensitivity)/specificity]	0.15 (0.07-0.30)
Diagnostic odds ratio [(sensitivity/(1-sensitivity))/{(1-specificity)/specificity}]	28.5 (9.41-85.6)
Error odds ratio [(sensitivity/(1-sensitivity))/{specificity/(1-specificity)}]	2.03 (1.46-3.34)
Youden J statistic [sensitivity+specificity-1]	0.67 (0.51-0.78)
Kappa value	0.66 (0.40-0.76)

CI = confidence interval.

consistent with previous studies [16], may be considered excellent and indicative of strong rank order stability of MDI-A scores. However, it has been argued that the calculation of correlation coefficients is not a sufficient method to test reliability and reproducibility because it is an index of correlation and not an index of agreement [17] and is more appropriately replaced by the ICC which shows the percentage of variance accounted for by the test-retest variance [18]. In the current study the ICC was also computed and, as might be expected, ICC was found to be excellent for the total score and the individual items of the MDI-A.

The relationship between the MDI-A and another well-validated measure of depression, the BDI, was examined to provide evidence of concurrent validity. As expected, we found a strong positive correlation between MDI-A and BDI scores. This finding is in line with previous research in other cultures, which cross-validated the MDI with other measures of depression, such as the Center for Epidemiological Studies-Depression and the Zung Depression Rating Scale, and reported high positive correlations [16].

Evidence for discriminant validity of the MDI-A was found, as indicated by the poor correlations with STAI-S and STAI-T which are better measures for anxiety-state and anxiety-trait respectively than measures for depression. Although anxiety and depression are related constructs, correlations were significantly higher between the STAI-S and STAI-T than between either of these measures and the MDI-A or BDI.

This study, however, has some possible limitations. The relatively small sample size is a common reason for both type I and type II errors. Another potential limitation is that the authors, test administrators and data analysts were not blind to the participants' diagnosis. Another factors to be considered is that the patient group comprised only individuals who met the diagnostic criteria

for major depressive disorder. The exclusion of patients with other diagnoses limits the variability of responses on the MDI-A and potentially limits the assessment of validity and test norming. A more heterogeneous sample would have more accurately tested this instrument as a screening tool. Other forms of reliability and validity could have been employed to strengthen the evidence of the applicability of the MDI-A. Finally, it should be noted that some colloquial words or phrases had to be added in the translation/adaptation process and that the study included people from only

one governorate of lower Egypt. This would limit the generalizability of the results to other populations in the Arab region.

Despite these limitations, our data indicate that the MDI-A had excellent reliability and an acceptable concurrent and discriminant validity. Future studies in a variety of situations are needed to strengthen and improve this study. It is hoped that the development of a valid and reliable version of the MDI in Arabic language will be a useful addition to both research and clinical settings.

Acknowledgements

The authors are grateful to Professor Per Bech, Department of Psychiatry, Frederiksberg General Hospital, Hillerød, Denmark, for permission to translate the MDI into Arabic. The help of Susan Søndergaard, Language Consultant, Mental Health Centre North Zealand, Hillerød, Denmark, in the translation process is acknowledged.

The authors have no biomedical financial interests or potential conflicts of interest to declare.

References

1. Kavasis I et al. Translation and validation of the Depression Outcomes Module (DOM) in Greece. *International Journal of Methods in Psychiatric Research*, 2005, 14:221-229.
2. Beck AT et al. Comparison of the Beck Depression inventories IA and II in psychiatric outpatients. *Journal of Personality Assessment*, 1996, 132:381-385.
3. Bech P et al. The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. *Journal of Affective Disorders*, 2001, 66:159-164.
4. Cuijpers P et al. Sensitivity and specificity of the Major Depression Inventory in outpatients. *BioMed Central Psychiatry*, 2007, 7:39.
5. Bech P, Wermuth L. Applicability and validity of the Major Depression Inventory in patients with Parkinson's disease. *Nordic Journal of Psychiatry*, 1998, 52:305-309.
6. Sartorius N, Kuyken W. Translation of health status instruments. In: Orley J, Kuyken W, eds. *Quality of life assessment: international perspective*. Berlin, Springer Verlag, 1994.
7. First MB et al. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, research version, non-patient edition (SCID-I/NP)*. New York, Biometrics Research Department, New York State Psychiatric Institute, 2002.
8. Beck AT et al. An inventory for measuring depression. *Archives of General Psychiatry*, 1961, 4:53-63.
9. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory. Twenty-five years of evaluation. *Clinical Psychology Review*, 1988, 8:77-100.
10. Shaw BF, Vallis TM, McCabe SB. The assessment of the severity and symptom patterns in depression. In: Beckham EE, Leber WR, eds. *Handbook of depression: Treatment, assessment, and research*. Homewood, Illinois, Dorsey Press, 1985.
11. Abdel-Khalek A. *Beck Depression Inventory: the Arabic version*. Cairo, Anglo-Egyptian Bookshop, 1996.
12. Spielberger CD, Gorsuch RL, Lushere RE, eds. *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California Consulting Psychologists Press, 1970.
13. Abdel-Khalek A. *State-Trait Anxiety Inventory: Arabic version*. Cairo, Anglo-Egyptian Bookshop, 1992.
14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing reliability. *Psychological Bulletin*, 1979, 86:420-428.
15. Fleiss J. *The design and analysis of clinical experiments*. New York, John Wiley and Sons, 1986.
16. Fountoulakis KN et al. Reliability, validity and psychometric properties of the Greek translation of the Major Depression Inventory. *BioMed Central Psychiatry*, 2003, 3:2.
17. Bartko JJ, Carpenter WT Jr. On the methods and theory of reliability. *Journal of Nervous and Mental Disease*, 1976, 163:307-317.
18. Streiner DL. Learning how to differ: agreement and reliability statistics in psychiatry. *Canadian Journal of Psychiatry*, 1995, 40:60-66.