International Journal of
Mental Health Systems

RESEARCH

Open Access

Development of a culturally sensitive Arabic version of the Mini International Neuropsychiatric Interview (M.I.N.I.-AR) and validation of the depression module

Carine Karnouk[1†], Kerem Böge[1†], Nico Lindheimer[1], Dana Churbaji[1], Shaymaa Abdelmagid[2], Sara Mohamad[1], Eric Hahn[1†] and Malek Bajbouj[1*†]

Abstract

Background: Arabic represents one of the most frequently spoken languages worldwide, especially among refugee populations. There is a pressing need for specialized diagnostic tools corresponding to the DSM-5 criteria in Modern Standard Arabic, which can be administered on Arabic speakers in the West and Arab region alike.

Objectives: To develop and validate the culturally-adapted version of the most recent M.I.N.I. 7.0.2 into Modern Standard Arabic—a form of Arabic commonly used across all Arab countries.

Methods: 102 participants were recruited between April 2019 to March 2020 at the Charité - Universitätsmedizin in Berlin. Symptoms were assessed with Arabic versions of rater-based and self-rated measures, including Mini International Neuropsychiatric Interview (M.I.N.I.), Patient Health Questionnaire (PHQ-9), and Harvard Trauma Questionnaire (HTQ). Arabic-speaking psychiatrists saw participants for diagnostic assessment.

Results: Cohen's kappa (κ) values were moderate for major depression, and slight for post-traumatic stress disorder, as well as generalized anxiety disorder. Moreover, kappa values indicated moderate agreement between M.I.N.I.-AR and PHQ-9 for depression, as well as HTQ for post-traumatic stress disorder, respectively.

Conclusion: The translated and culturally adapted version of the M.I.N.I. addresses an existing need for a reliable, efficient, and effective comprehensive diagnostic tool using the most recent DSM-5 criteria in Modern Standard Arabic (MSA). Based on the obtained results, only a validation of the depression module (Module A) of the M.I.N.I-AR was possible. Study outcomes also show evidence for the validation of Module H covering Post-Traumatic Stress Disorder. Potential valuable contributions can be extended to this translation and validation.

Keywords: Validation, MINI, Psychiatric diagnosis, Arabic, Cultural adaptation, Assessment, Arabic-speakers, Refugee

*Correspondence: malek.bajbouj@charite.de
†Carine Karnouk, Kerem Böge, Eric Hahn, and Malek Bajbouj contributed equally to this work
[1] Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin, Campus Benjamin Franklin, Berlin, Germany, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Campus Benjamin Franklin, Hindenburgdamm 30, 12203 Berlin, Germany
Full list of author information is available at the end of the article

Introduction

There are as many as 274 million native Arabic speakers worldwide, making Arabic the 6th most spoken language [1]. Furthermore, 25 countries consider Arabic their official or co-official national language. Nowadays, Arabic is not only spoken in indigenous settings, but due to migration and refugee resettlements, Arabic-speakers now live dispersed globally [2, 3]. According to the German

Karnouk *et al. Int J Ment Health Syst* (2021) 15:24

Page 2 of 10

Federal Office for Migration and Refugees [4], Arabic is the most frequently spoken native language within Germany's refugee population. In 2018, it was also reported that Arabic-speaking refugees had shown an increase in the prevalence of mental disorders [3] and are thus in need of accurate and efficient, culturally sensitive diagnostic assessments and treatment plans [2]. Furthermore, in Arabic-speaking countries, the burdens of mental health have been estimated to be much higher than in Western counterparts [5, 6]. Hence, Arabic-speakers face several disadvantages not only in their home countries but also in host countries and humanitarian aid settings [2, 7, 8]. Some of the challenges that they experience include lack of available effective and efficient culturally sensitive mental health services, cultural misunderstandings, stigma, and language barriers [2, 7, 8]. The aforementioned barriers have led to poor diagnoses and treatment choices.

Despite the high demand for specialized services for Arabic-speakers, a considerably large treatment gap exists [5, 9]. Furthermore, large-scale epidemiological studies and national data are scarce in the Arab region [10], making it challenging to estimate psychiatric prevalence rates and needs. In a study summarizing mental health services across the Arab region, it was found that there were less than 1 to 5 psychiatrists for every 100,000 inhabitants in 22 Arab countries [9]. Moreover, a comprehensive assessment using multiple search engines, found that in about 30 years (from 1966 to 1996), only a total of 1058 research papers were produced addressing the topic of mental health in the Arab world [10]. However, in the past decade, research contributions in the Arab world have reportedly increased by almost 160%, compared with only 57% for other regions. Although this number represents exponential growth (particularly in Egypt, Saudi Arabia and Lebanon), in total the number of research outputs is still comparably lower than the rest of the world [11]. The dearth of literature may explain the scarcity in available professional resources, assessment tools and specialized expertise in the region [5, 10], leading some Arab psychiatrists/psychologists to seek training in western contexts, thereby ignoring essential cultural pillars [8, 10]. As a result, over the years many unvalidated assessments and treatment methods have been adopted and used with Arab populations yielding to 'culturally incongruent' applications [2, 10]. However, in the last decade, significant progress has been made to address these issues in Arab and Western contexts resulting in increased research and hybrid mental health interventions taking elements from multiples cultures into consideration [7, 8, 10].

Clinical structured interviews were developed for diagnostic precision, speed and accuracy and carry several advantages. Because the questions are precise and elicit only a limited number of responses, these assessments can largely be administered by non-specialized interviewers, leading to high inter-rater reliability. Special algorithms are built to distinguish between clinically significant symptoms from regular stressors—a feature that can be used to ensure homogenous groups in clinical trials. Among the most commonly used clinical structured interviews internationally, are the Mini International Neuropsychiatric Interview (MINI) [12], Structured Clinical Interview (SCID), Present State Examination (PSE), Diagnostic Interview Schedule (DIS) and the Composite International Diagnostic Interview (CIDI) [13, 14]. From these interviews, only the DIS [15], the PSE-10 [16] and the CIDI [17], exist in Modern Standard Arabic, in addition to a Moroccan dialect validated version of the M.I.N.I. [18].

The M.I.N.I. was developed by Sheehan et al. [14] to explore some psychiatric disorders according to the DSM and ICD diagnostic criteria. Every few years, the M.I.N.I. is updated and validated versions in other languages also become available [14]. Initially, the M.I.N.I. was designed to meet the need for a short yet valid and reliable psychiatric interview for multicenter clinical trials and epidemiological studies. However, in recent years, the assessment tool is also being used in humanitarian aid and global health settings [7]. The M.I.N.I. is particularly attractive since it offers a fee waiver for researchers and clinicians who will use the assessment to assist refugees or victims of terrorism (for more information see https://harmresearch.org/index.php/mini-international-neuropsychiatric-interview-mini/). This stipulation may be more relevant now than ever, as refugee groups resettle worldwide.

The Moroccan version of the M.I.N.I. developed by Kadri et al. [18] corresponds to the DSM-IV criteria and was compared with expert diagnoses. This version uses the colloquial Moroccan dialect, which is quite different from Modern Standard Arabic and also includes French. The M.I.N.I. has been widely used in several studies in the Arab world, however, the majority of these studies developed their translations and focused on specific modules, such as post-partum depression [19], post-traumatic stress disorder [20], alcohol abuse and dependence [21] and schizophrenia [22, 23]. To the authors' knowledge, there is no fully translated, culturally adapted and validated version of the M.I.N.I. available in Modern Standard Arabic (MSA) corresponding to the latest DSM-5 criteria.

This study aims to develop and validate a culturally-adapted version of the M.I.N.I. 7.0.2 in Modern Standard Arabic. Although various Arabic dialects exist, MSA—a standardized, formal version of the language,

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 3 of 10

most commonly used in schools, the press and other official contexts—is widely understood by all native Arabic speakers.

Methods

Procedure and participants

The current cross-sectional validation study of the Arabic-version of the Mini International Neuropsychiatric Interview (M.I.N.I.-AR) was conducted to address the existing gap in available validated diagnostic assessment tools in Modern Standard Arabic language. Therefore, this adaption was developed in response to a need for the M.I.N.I.-AR to be used in a large German multi-center trial for Arabic speaking refugees in Germany (MEHIRA: Mental Health in Refugees and Asylum Seekers) [7]. The MEHIRA project depicts a multicenter randomized control trial investigating a stepped care and collaborative model (SCCM) for refugees and asylum seekers across eight university cities in Germany. All participants for the current trial were recruited through the similar recruitment channels of the MEHIRA network, which include regionally heterogeneous allocation paths, such as general practitioners, central clearing clinics, residential care settings, and social agencies in Berlin, Germany. Assessments took place either at the Charité - Universitätmedizin Berlin central clearing clinic, a specialized first-contact outpatient facility for refugees and asylum seekers in Berlin, or the Arabic-speaking outpatient clinic of the Department of Psychiatry and Psychotherapy, Campus Benjamin Franklin—both institutions led by the Charité - Universitätsmedizin in Berlin, Germany.

A sample of 102 participants was recruited between April 2019 to March 2020. Inclusion criteria were defined as age between 18 and 75 years and native Arabic speakers. Initially, two interviewers received in-depth structured training and supervision by two licensed psychologists to ensure consistency in the administration of the interview process. In the next step, participants received an information sheet about the study aims and were asked for their interest in participation. Throughout the whole study process, participants were encouraged to ask any questions that remain unclear. Participants who were interested and willing to participate gave their signed informed consent before any trial-related procedures were conducted. Study participation was voluntary, and no monetary compensation was provided. In the next step, participant symptomatology was assessed by using rater-based and self-rated measures, including the Mini International Neuropsychiatric Interview (M.I.N.I.), the Patient Health Questionnaire (PHQ-9), and the Harvard Trauma Questionnaire (HTQ). All assessments were conducted in Arabic. In parallel, participants were also seen by a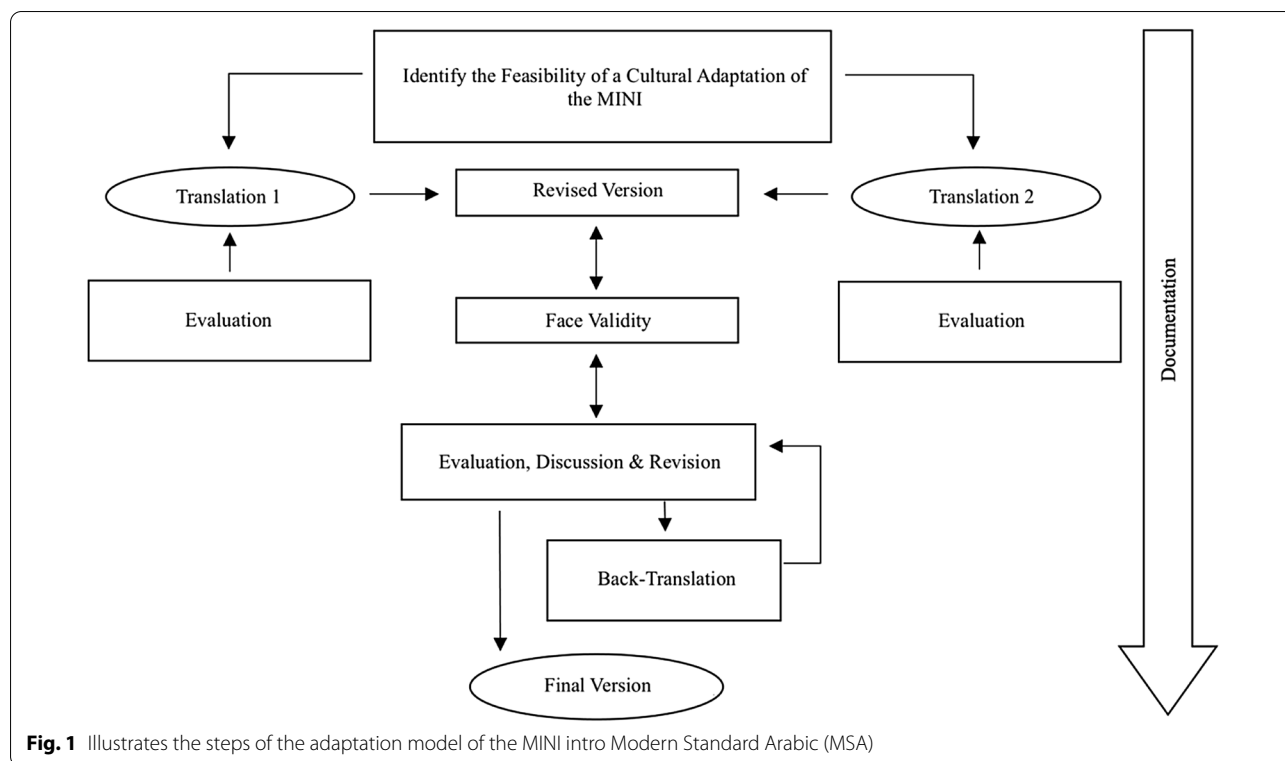n Arabic-speaking licensed psychiatrist for a diagnostic assessment and consultation. Three separate psychiatrists were available in rotation at the outpatient clinic. Both interviewers and the psychiatrist were blind to the diagnosis of each other. The whole assessment procedure took between 60 and 90 min. The study was approved by the ethics committee of Charité - Universitätsmedizin Berlin, Germany, and is in line with the Declaration of Helsinki.

The Mini International Neuropsychiatric Interview (M.I.N.I.)

For the current study, the M.I.N.I 7.0.2. validated English version was used and translated into Modern Standard Arabic (MSA, M.I.N.I.-AR; Additional file 1). The M.I.N.I. is a short (needing between 15 and 30 min for completion), structured diagnostic interview compatible with the DSM-5 and ICD-10 criteria. It was specifically designed for implementation in clinical practice and research in psychiatric-, as well as primary health care settings. The M.I.N.I. includes 130 questions with "yes" and "no" answer options examining 16 axis-I DSM-5 disorders as well as one personality disorder. Each of the 16 modules starts with a screening question to exclude the diagnoses and possibly skip the module accordingly if answered negatively or explore symptoms severity when responded positively. Several validation studies [12, 24] have demonstrated excellent interrater and test reliabilities of the M.I.N.I. Furthermore, moderate validity with both the extensive Composite International Diagnostic Interview (CIDI) [12, 14] and the Structured Clinical Interview for DSM-4 (SCID) [12, 14] have been exhibited. Through a previous trial by our research group with a pilot sample size (N=20), initial validity has been shown for the translated M.I.N.I.-AR [25].

Cultural adaptation and translation

In a pilot trial published by our research group [25], a translation and cultural/linguistic adaptation of the M.I.N.I. 5.0.0 version was carried out. As a follow-up project, the validation of the most recent M.I.N.I. 7.0.2 version, with the same constellation of a multilingual, interdisciplinary team, including five psychologists and three psychiatrists, collaborated on a multilevel adaption process for the translation into the Arabic language (M.I.N.I.-AR). Members of the research group who were involved in the translation process are native Arabic speakers from different Arab nations. They have extensive experience in the mental health sector, as well as board knowledge and familiarity with the cultures of the original (English) and targeted language (Arabic). According to the World Health Organization guidelines [26], for the translation and adaptation process of instruments and aligned with Brislin [27], the steps of the adaption model are illustrated in Fig. 1.

Fig. 1 Illustrates the steps of the adaptation model of the MINI intro Modern Standard Arabic (MSA)

In the first step, two translators (psychologists) evaluated the M.I.N.I. concerning its feasibility and practicability for a culturally-sensitive adaptation and translation. Secondly, both independently translated, assessed and revised all items. Afterward, potentially challenging and unclear items were discussed with a further psychologist (native English- and Arabic speaker) for accuracy and cultural relevance, and the translations were merged into one main revised version. To ensure face validity one trained interviewer according to the criteria by Lecrubier et al. [24], revised the version, and subsequently conducted it together with a licensed psychiatrist on Arabic speaking refugees as recommended by Bannigan and Watson [28].

Consequently, any unclear, problematic and challenging phrasing and linguistic expressions mentioned by the participants and the research staff were adjusted and adapted under special considerations of the Arab culture and cultural idioms within an iterative process of various feedback and revision loops (see example Table 1). Parallelly, a back-translation was performed into English by a blind, independent and bilingual psychologist to the original MINI-7.0 version according to international guidelines [26, 27]. In a final evaluation and revision, five psychologists and three psychiatrists discussed and adjusted outcomes until an agreement for a final version was reached. For a more detailed description of

the cultural and linguistic adaptation and translation, see Churbaji et al. [25].

Questionnaires

The Patient Health Questionnaire-9 (PHQ-9) is a short and concise self-rating screening tool to assess the severity of depressive symptoms. It consists of nine items that cover all symptoms of a major depressive episode according to DSM-4 criteria [29]. Responses are given by participants on a four-point Likert scale (0 "not at all" to 3 "almost every day"), indicating their symptom severity for the past 2 weeks. Total scores can range from 0 to 27, with higher values indicating symptom severity. An overall score > 10 indicates the presence of depression. Cut-off values for symptoms severity are as following: minimal (2–9), mild (10–14), moderate (15–19), and severe ($\geq$ 20). For the Arabic translation of the PHQ-9, a discriminant, factorial, and convergent validity and high reliability ($0.86 \leq \alpha \leq 0.88$) have been shown [30].

The Harvard Trauma Questionnaire (HTQ) is a well-established self-reported screening instrument for trauma, torture, and post-traumatic stress disorder (PTSD) symptoms [31]. It covers three parts with 42 items, an open-ended question part, and 16 items, respectively. For the current study, only the third part was included to assess the severity of PTSD symptoms. Participants report their experiences with trauma

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 5 of 10

Table 1 Examples of the cultural adaptation of the MINI

Question MINI	Translation 1	Final translation	Type of equivalence	Explanation
Have you been consistently depressed or down most of the day, nearly every day, for the past two weeks?	‏هل تشعر باستمرار بالاكتئاب أو الإحباط طوال اليوم، معظم الأيام، تقريباً خلال الأسبوعين الماضيين؟‏	‏هل تشعر باستمرار بالحزن أو الاكتئاب معظم اليوم، تقريباً كل يوم، خلال الأسبوعين الماضيين؟‏	Vocabulary equivalence	"down" ↔ "sad". The term "down" can't be used to describe emotions in the Arabic language
Did you have trouble sleeping nearly every night (difficulty falling asleep, waking up in the middle of the night, early morning wakening or sleeping excessively)?	‏هل تعاني من مشاكل في النوم كل ليلة تقريباً (صعوبة في النوم، الاستيقاظ في منتصف الليل، الاستيقاظ في الصباح الباكر أو النوم بكثرة)؟‏	‏هل تعاني من مشاكل في النوم كل ليلة تقريباً (صعوبة في النوم، الاستيقاظ في منتصف الليل، الاستيقاظ في الصباح الباكر أو النوم بكثرة)؟‏	Idiomatic equivalence	"falling asleep" ↔ "go to sleep". The term "falling asleep = ‏نوّم‏" is used in the Arabic language for a nap
In the past 12 months, have you had 3 or more alcoholic drinks within a 3-h period on 3 or more occasions?	‏خلال الـ ١٢ شهراً الماضية، هل تناولت ثلاث كاسات كحول أو أكثر في فترة ثلاث ساعات ٣ مرات أو أكثر؟‏	‏خلال الـ ١٢ شهراً الماضية، هل تناولت ثلاث كاسات كحول أو أكثر في فترة ثلاث ساعات ٣ مرات أو أكثر؟‏	Experiential equivalence is not given	"alcoholic drinks" ↔ "alcoholic pots". The quantity was not clearly understood. Even after adjustment, questions were asked about the size and the strength of the alcoholic drinks

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 6 of 10

symptoms such as "feeling detached or withdrawn from people", "difficulty concentrating", or "trouble sleeping" on a four-point Likert scale (1 "not at all" to 4 "extreme") for each of the 16 items. A total score is calculated by dividing the sum of scores of items 1–16 by the number of items answered, with higher scores indicating an ascending level of PTSD symptom severity. The value can be between 1 and 4 with 2.5 as a cut-off value for a present PTSD, according to DSM-4 [31]. The scale has been validated in numerous cultures and languages, including Arabic, and has shown good psychometric properties [32, 33].

Statistical analysis

All data was collected and stored pseudonymized in a spreadsheet using Package for the Social Science (SPSS) 25, MacOS-X. Statistical analyses were set at an exploratory significance level of $p < 0.5$. Descriptive measures of the sample are summarized as means and standard deviations (SD) for continuous measures and as frequencies and percentages for categorical variables. The criterion validity of the M.I.N.I.-AR was tested by comparing each M.I.N.I. diagnosis with the clinical diagnosis of a blind psychiatrist and its concordance with two screening tools namely the Patient Health Questionnaire 9 (PHQ-9) [29] for depressive symptoms (cut-off score set at 10), and the Harvard Trauma Questionnaire (HTQ) [34] (cut-off score set at 2.5) to detect the presence of post-traumatic stress symptoms. Numerous studies have highlighted that the assessment of criterion validity through the use of clinical diagnoses is a well-established method [13, 18]. Cohen's kappa (κ) was used as a statistically adjusted measure of concordance between two measures [35]. It has been recommended that only for diagnoses with a frequency of 5% or more kappa values are calculated [13]. According to Landis and Koch [36] the following values were used to interpret significant results: < 0 no agreement, 0–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, 0.81–1.0 perfect. Additionally, statistics for specificity, sensitivity, negative (NPV) and positive predictive values (PPV), and the number of false-positive and false-negative diagnoses were calculated.

Results

Clinical and sociodemographic characteristics

In total, the data of 102 Arabic-speaking refugee outpatients, 45 males, and 57 females, were analyzed. The mean age was 35.29 years (SD = 9.66; range 19–61). The vast majority of the sample named Syria as their country of origin (81%), followed by Iraq (7.8%) and Lebanon (4.5%). All sociodemographic characteristics assessed are illustrated in Table 2.

Table 2 Sociodemographic characteristics of the study sample

Sociodemographic data	$N = 102$
Gender	
Male	45 (44.1%)
Female	57 (55.9%)
Age in years M ± SD 35.29 ± 9.66*	$n = 97$
19–30	35 (34.3%)
31–40	35 (34.3%)
41–50	19 (18.6%)
51–61	8 (7.8%)
Country of origin	
Syria	81 (79.4%)
Iraq	8 (7.8%)
Lebanon	5 (4.9%)
Egypt	4 (3.9%)
Palestine	3 (2.9%)
Morocco	1 (1.0%)

Across all measures used, only major depressive episode or disorder (MDE) (M.I.N.I.: 83.3%, Expert: 61.5%, PHQ-9: 74.5%), posttraumatic stress disorder (PTSD) (M.I.N.I.: 36.3%, Expert: 14.6%, HTQ: 52.0%), and generalized anxiety disorder (M.I.N.I.: 36.6%, Expert: 5.2%) had a occurance of < 5% in the M.I.N.I. screening as well as expert diagnoses. Detailed prevalence estimates for each disorder of the present sample by at least one of the measures assessed is depicted in Table 3.

M.I.N.I.-AR concordance with clinical diagnoses

A total of 96 individuals received a diagnosis from both the M.I.N.I.-AR and a clinical expert. Since only major depressive episode, post-traumatic stress disorder, and generalized anxiety disorders had a prevalence of < 5% in the M.I.N.I. screening and expert diagnoses, all other diagnoses were excluded from the validity assessment. Generally, kappa values were moderate for major depression episodes/disorders (0.54) and slight for post-traumatic stress disorders (0.2), and generalized anxiety disorder (0.12). All results for the three disorders are depicted in Table 4.

M.I.N.I.-AR concordance with the PHQ-9 and HTQ

102 participants were administered the M.I.N.I.-AR and filled self-rating measures PHQ-9 and HTQ. Kappa values indicated moderate agreement between M.I.N.I.-AR and PHQ-9 (0.58) as well as HTQ (0.53), respectively. Detailed results are summarized in Table 5.

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 7 of 10

Table 3 Prevalence estimates, depending on the measure used

Prevalence estimates	M.I.N.I N = 102	Expert N = 96	PHQ/HTQ N = 102
A. Major depressive episode/ disorder	78 (76.5%)*	59 (61.5%)	76 (74.5%)
C. Manic episode/bipolar disorder	2 (2.0%)	1 (1.0%)	
D. Panic disorder	34 (33.3%)	3 (3.1%)	
E. Agoraphobia	12 (11.8%)	0 (0.0%)	
F. Social anxiety	14 (13.7%)	4 (4.2%)	
G. Obsessive–compulsive disorder	16 (15.7%)	2 (2.1%)	
H. Post-traumatic stress disorder	37 (36.3%)	13 (13.5%)	53 (52.0%)
I. Alcohol use disorder	3 (2.9%)	3 (3.1%)	
J. Substance use disorder	3 (2.9%)	3 (3.1%)	
K. Psychotic disorder	0 (0.0%)	1 (1.0%)	
L. Anorexia nervosa	1 (1.0%)	0 (0.0%)	
M. Bulimia nervosa	1 (1.0%)	0 (0.0%)	
MB. Binge-eating disorder	0 (0.0%)	0 (0.0%)	
N. Generalized anxiety disorder	11 (10.8%)	9 (9.4%)	
P. Antisocial PS	7 (6.9%)	0 (0.0%)	
Adjustment disorder		12 (12.5%)	
Somatoform disorders		3 (3.1%)	
Dissociative disorders		3 (3.1%)	
Sexual dysfunctions		1 (1.0%)	
Emotionally unstable PS		1 (1.0%)	
Hyperkinetic disorders		1 (1.0%)	
Nonorganic sleep disorders		1 (1.0%)	
Histrionic PS		1 (1.0%)	
Pathological gambling		1 (1.0%)	

*Excluding 7 past episodes. Multiple diagnoses per individual are possible

M.I.N.I. Mini International Neuropsychiatric Interview, *Expert* clinical expert diagnoses, *PHQ* Patient Health Questionnaire 9, *HTQ* Harvard Trauma Questionnaire, *PS* personality disorder

Discussion

The overall aim of the current study was to validate a translated and culturally adapted version of the most recent M.I.N.I. 7.0.2 into Modern Standard Arabic—a form of Arabic commonly used across all Arab countries in official settings and the press. Although the obtained results did not show evidence for the validation of all the modules within the instrument, it was possible to validate the depression module (Module A). Results showed moderate agreements between Module A of the M.I.N.I.-AR 7.0.2 compared with expert diagnosis and the PHQ. Also noteworthy to reveal were the moderate to slight agreements between M.I.N.I.-AR, HTQ and expert diagnosis concerning post-traumatic stress disorder. The M.I.N.I.-AR 7.0.2 was translated and culturally-adapted to address a diagnostic gap in a large German research trial for the psychological treatment of refugees of Arab descent [7]. Together with our pilot trial [25], this Arabic version of the M.I.N.I. can contribute to the existing literature and be used in mental health care settings serving Arabic-speaking populations as a whole. The adapted instrument is relevant to both Arab and Western contexts alike. Nonetheless, further studies and extensions may use this carefully adapted translation and compare it with a gold standard structured interview that may offer more robust results.

Our study outcomes are in line with a similar validation study using the same method [13] and our initial pilot study [25]. The original authors of the M.I.N.I. also reported a rather poor agreement between expert diagnoses and M.I.N.I. diagnoses in their validation studies [14]. This is not surprising, since the M.I.N.I. was originally designed to be oversensitive. As reported by Sheehan et al. [14], over-diagnosing cases may be less harmful than missing a case. One explanation could be

Table 4 Concordance of the M.I.N.I.-AR with clinical diagnoses

Disorder		Expert diagnosis		Kappa (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
	M.I.N.I.	−	+					
	−	TN	FN					
	+	FP	TP					
Major depressive episode/disorder		19	1	0.54 (0.38 to 0.71)	98.31 (90.91 to 99.96)	51.35 (34.40 to 68.08)	74.32 (67.79 to 81.80)	95.00 (72.63 to 99.27)
		18	58					
Post-traumatic stress disorder		55	4	0.20 (0.04 to 0.37)	69.23 (38.57 to 90.91)	66.27 (55.05 to 76.28)	24.32 (16.71 to 33.99)	93.22 (85.71 to 96.93)
		28	9					
Generalized anxiety disorder		79	7	0.12 (− 0.13 to 0.39)	22.22 (2.81 to 60.01)	90.86 (82.68 to 95.95)	20.00 (5.87 to 50.07)	91.86 (88.78 to 94.15)
		8	2					

N = 96

M.I.N.I. Mini International Neuropsychiatric Interview, *Expert* clinical expert diagnoses, *CI* confidence interval, *TP* true positive, *FP* false positive, *FN* false negative, *TN* true negative, *PPV* positive predictive value, *NPV* negative predictive value

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 8 of 10

Table 5 Concordance of the M.I.N.I.-AR with the PHQ-9 and the HTQ

Disorder	M.I.N.I.	PHQ-9/HTQ		Kappa (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
		−	+					
	−	TN	FN					
	+	FP	TP					
Major depressive episode/disorder		17	7	0.58 (0.39 to 0.76)	90.79 (81.94 to 96.22)	65.38 (44.33 to 82.79)	88.46 (81.81 to 92.89)	70.83 (53.20 to 83.84)
		9	69					
Post-traumatic stress disorder		45	20	0.53 (0.38 to 0.69)	62.26 (47.89 to 75.21)	91.84 (80.40 to 97.73)	89.19 (75.91 to 95.57)	69.23 (61.19 to 76.25)
		4	33					

$N = 102$

M.I.N.I. Mini International Neuropsychiatric Interview, *Expert* clinical expert diagnoses, *CI* confidence interval, *TP* true positive, *FP* false positive, *FN* false negative, *TN* true negative, *PPV* positive predictive value, *NPV* negative predictive value

that experts may not be able to review all the information available and thus overlook diagnoses [37]. In contrast, the M.I.N.I. is a particularly comprehensive tool offering fast, efficient and accurate diagnoses. This is a particular asset of the M.I.N.I. since it can be administered by non-specialists leading to resource-saving and cost-effectiveness—both highly relevant benefits, especially in the case of Europe's refugee resettlements and the scarcity of resources and qualified personnel in the Arab world.

Recent studies have highlighted that patients from a Muslim cultural background may experience feelings of discomfort when alone with a therapist from the opposite gender [38], sometimes resulting in an inaccurate representation of symptoms during psychiatric evaluation. In the present study, gender may have played a role in participant responses. Matching gender between psychiatrists and patients was not always possible in the expert interviews, whereas there was more flexibility in the administration of the M.I.N.I.. Since the M.I.N.I. was designed to be administered by non-specialists, it provides an efficient solution to the shortage of available specialized professional care, who sometimes lack the language skills and cultural competence training needed in diagnosing Arabic-speaking populations. Taken together, these reports may explain the slight to moderate Kappa values when comparing the M.I.N.I.-AR and expert diagnoses. Nonetheless, the specific effect of matching gender was not within the scope of this paper, however, it may be interesting to assess this in future validation studies.

The present study has several strengths, especially when considering the limited availability of diagnostic tools and research targeting this specific population. A key strength of the present study is the careful translation and cultural adaptation process of the M.I.N.I. 7.0.2 into Modern Standard Arabic. As mentioned above, Modern Standard Arabic is particularly useful since it is a form of Arabic that is pluricentric—a major practical advantage

for Arabic speakers and clinicians worldwide. Language and cultural barriers are one of the factors contributing to the underrepresentation of minorities in clinical settings and research samples [39]. This instrument can thus be used to increase the representation of minorities from the Arab region in clinical practice and research. The translation and adaptation process of the M.I.N.I. was based on experiences from our pilot study [25] and included the same multilingual and interdisciplinary team of psychologists and psychiatrists, each with their unique set of expertise. The team included Arabic-speakers from different Arab countries. The translation and adaptation followed the WHO guidelines [26] and were in line with the adaptation model developed by Brislin [27]. Furthermore, the use of culturally appropriate stories or metaphors to emphasize key concepts helps keep patients engaged and can be used in the diagnosis and different stages of the therapeutic encounter [40]. Sometimes, patients are unable to name an illness to describe their condition—in some languages, words used to name emotions may not correspond to those used in psychiatric settings in the West. For example, using words such as feeling "down" or "high" to describe an emotion does not exist in Arabic. Careful consideration was given to the meaning of the translation to ensure exact understanding. Therefore, by including culturally appropriate words to describe emotions, this instrument can improve the quality of diagnosis and therapeutic encounters when applied with Arabic-speakers and refugee communities in the West [38] and in the Arab region. The M.I.N.I-AR also contains a comprehensive table of drugs that are commonly used in the Arab region including their local street names.

The findings of this study have to be seen in the light of several limitations. To conduct a robust validation, a gold standard structured interview coupled with a larger sample size and a more clinically diverse sample is needed. This was not possible due a scarcity of

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 9 of 10

resources (clinic capacity), and the vulnerability of our sample (difficultry concentrating and fatigue), therefore valuable future contributions can be extended to this validation. Furthermore, study results show a high count of adjustment disorders in expert diagnoses, but no possibility to diagnose this with the M.I.N.I. Indeed, this is a weakness of the M.I.N.I. In many settings, the diagnosis of adjustment disorder seems less stigmatizing and more easily acceptable [41] and may thus be especially appropriate for individuals from the Arab world, where mental health stigma plays an important role [42]. Nonetheless, similar symptoms are present in depression and PTSD, so they are accounted for. A further limitation is a comparison of HTQ and PHQ (DSM-IV) with the M.I.N.I. trauma section and depression sections (DSM-5). Even though there was a considerable amount of changes from the DSM-IV to 5 in the definition of trauma [43], and rather small changes in the depression section [44], the questions in the M.I.N.I. versions remained mostly the same and should thus be appropriate for the comparison with questionnaires using the DSM-IV as a basis. This may also be a consideration for future research. Moreover, study results also revealed a low prevalence of general disorders. In our sample, the most prevalent diagnoses were depression, PTSD, and anxiety; therefore, reliable kappa values could not be calculated for many disorders. Yet, the proportion of disorders in the current study mirrors the prevalence of the most reported disorders in the refugee community [45], reflecting ecological validity. The M.I.N.I-AR was validated on a sample made up of mostly young Syrians, and Arab patients living in the West may differ from those living in their country of origin [38]. Therefore, for generalizability, future directions for this research may target Arab populations with a more diverse representation of psychiatric disorders.

In conclusion, our translated and culturally-adapted M.I.N.I-AR addresses an existing gap and need for a reliable, efficient, and effective diagnostic tool that can be used in both clinical and research settings involving Arabic speakers worldwide. Based on the obtained results, we were only able to validate module A of the M.I.N.I-AR covering depression and show some evidence in support of the validation of Module H covering Post-Traumatic Stress Disorder. Since no other diagnostic tool corresponding to the DSM-5 criteria is currently available in Modern Standard Arabic, the present translation and cultural adaptation particularly useful for Western host countries and humanitarian aid settings that are unable to find an alternative. Although the instrument was validated on a sample of Syrian refugees, it was developed in Modern Standard Arabic, leading to its potential usefulness in the entire Arab region.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13033-021-00447-1.

Additional file 1. M.I.N.I.-AR.

Authors' contributions
CK: study concept and design, management, implementation, translation, data collection, data interpretation and preparation of manuscript. KB: study design, methodology, management, data analysis, data interpretation and preparation of manuscript. NL: data analysis and contribution in preparation of manuscript. DC: translation, data collection and contribution in preparation of manuscript. SA: translation and review of manuscript. EH: study design, supervision, editing and review of manuscript. MB: study design, supervision, management, editing and review of manuscript. All authors read and approved the final manuscript.

Availability of data and materials
The dataset used and analysed in the current study is available upon request from the corresponding author.

Declarations

Ethics approval and consent to participate
Written consent was collected from all of the study participants. The study was approved by the ethics committee of Charité - Universitätsmedizin Berlin, Germany, and is in line with the Declaration of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin, Campus Benjamin Franklin, Berlin, Germany, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Campus Benjamin Franklin, Hindenburgdamm 30, 12203 Berlin, Germany. [2] Balance Psychology Center, Cairo, Egypt.

References
1. Eberhard DM, Simons GF, Fennig CD. Ethnologue: languages of Africa and Europe; languages of the Americas and the Pacific; languages of Asia. 23rd ed. Dallas: SIL International; 2020.
2. Bajbouj M, Alabdullah J, Ahmad S, Schidem S, Zellmann H, Schneider F, *et al*. Psychosoziale Versorgung von Flüchtlingen in Deutschland. Nervenarzt. 2018;89(1):1–7.

Karnouk *et al. Int J Ment Health Syst*   (2021) 15:24

Page 10 of 10

3. Georgiadou E, Zbidat A, Schmitt GM, Erim Y. Prevalence of mental distress among Syrian refugees with residence permission in Germany: a registry-based study. Front Psychiatry. 2018;9:393.

4. Bundesamt für Migration und Flüchtlinge. Aktuelle Zahlen zu Asyl. 2018.

5. Pocock L. Mental health issues in the Middle East: an overview. Middle East J Psychiatry Alzheimers. 2017;84(4221):1–6.

6. Karam EG, Mneimneh ZN, Karam AN, Fayyad JA, Nasser SC, Chatterji S, *et al*. Prevalence and treatment of mental disorders in Lebanon: a national epidemiological survey. Lancet. 2006;367(9515):1000–6.

7. Böge K, Karnouk C, Hahn E, Schneider F, Habel U, Banaschewski T, *et al*. Mental health in refugees and asylum seekers (MEHIRA): study design and methodology of a prospective multicentre randomized controlled trail investigating the effects of a stepped and collaborative care model. Eur Arch Psychiatry Clin Neurosci. 2020;270(1):95–106.

8. Karnouk C, Böge K, Bajbouj M, Hahn E, Schweininger S, Strasser J. Psychotherapy in Jordan: an investigation of the host-and Syrian refugee community's perspectives. Front Psychiatry. 2019;10:556.

9. Okasha A, Karam E, Okasha T. Mental health services in the Arab world. World Psychiatry. 2012;11(1):52–4.

10. Jaalouk D, Okasha A, Salamoun MM, Karam EG. Mental health research in the Arab world. Soc Psychiatry Psychiatr Epidemiol. 2012;47(11):1727–31.

11. Zeinoun P, Akl EA, Maalouf FT, Meho LI. The Arab region's contribution to global mental health research (2009–2018): a bibliometric analysis. Front Psychiatry. 2020;11:182.

12. Sheehan D, Lecrubier Y, Sheehan KH, Janavs J, Weiller E, Keskiner A, *et al*. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. Eur Psychiatry. 1997;12(5):232–41.

13. Otsubo T, Tanaka K, Koda R, Shinoda J, Sano N, Tanaka S, *et al*. Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. Psychiatry Clin Neurosci. 2005;59(5):517–26.

14. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, *et al*. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998;59(20):22–33.

15. Karam E, Barakeh M, Karam AN, El-Khouri N. The Arabic diagnostic interview schedule. Rev Med Libanaise. 1991;3:28–30.

16. Sabry N. Inter-rater reliability of the short Arabic form of the Present State Examination (PSE-10). Egypt J Psychiatry. 2009;29(2):45–54.

17. Karam E, Yabroudi P, Nasser-Karam A, Mansour C, Sabah S, Al AR. The Arabic composite international diagnostic interview (CIDI). Arab J Psychiatry. 1995;6:19–29.

18. Kadri N, Agoub M, El Gnaoui S, Alami KM, Hergueta T, Moussaoui D. Moroccan colloquial Arabic version of the Mini International Neuropsychiatric Interview (MINI): qualitative and quantitative validation. Eur Psychiatry. 2005;20(2):193–5.

19. Hamdan A, Tamim H. The relationship between postpartum depression and breastfeeding. Int J Psychiatry Med. 2012;43(3):243–59.

20. Kazour F, Zahreddine NR, Maragel MG, Almustafa MA, Soufia M, Haddad R, *et al*. Post-traumatic stress disorder in a sample of Syrian refugees in Lebanon. Compr Psychiatry. 2017;72:41–7.

21. Yazbek J-C, Haddad R, Bou Khalil R, Hlais S, Rizk GA, Rohayem J, *et al*. Prevalence and correlates of alcohol abuse and dependence in Lebanon: results from the Lebanese epidemiologic survey on alcohol (LESA). J Addict Dis. 2014;33(3):221–33.

22. Chamali R, Ghuloum S, Sheehan DV, Mahfoud Z, Yehya A, Opler MG, et al. Cross-validation of the Arabic Mini International Neuropsychiatric Interview, module K, for diagnosis of Schizophrenia and the Arabic positive and negative syndrome scale. J Psychopathol Behav Assess. 2020;42(1):139–49. https://doi.org/10.1007/s10862-019-09759-6.

23. Yehya A, Ghuloum S, Mahfoud Z, Opler M, Khan A, Hammoudeh S, *et al*. Validity and reliability of the Arabic version of the positive and negative syndrome scale. Psychopathology. 2016;49(3):181–7.

24. Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan KH, *et al*. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry. 1997;12(5):224–31.

25. Churbaji D, Lindheimer N, Schilz L, Böge K, Abdelmagid S, Rayes D, *et al*. Entwicklung einer Kultursensiblen Version des Mini-International Neuropsychiatric Interview (MINI) in hocharabischer Sprache. Fortschritte der Neurologie·Psychiatrie. 2020;88(02):95–104.

26. World Health Organization. Process of translation and adaptation of instruments. Geneva: World Health Organization; 2019.

27. Brislin RW. Back-translation for cross-cultural research. J Cross Cult Psychol. 1970;1(3):185–216.

28. Bannigan K, Watson R. Reliability and validity in a nutshell. J Clin Nurs. 2009;18(23):3237–43.

29. Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry. 2010;32(4):345–59.

30. AlHadi AN, AlAteeq DA, Al-Sharif E, Bawazeer HM, Alanazi H, AlShomrani AT, *et al*. An arabic translation, reliability, and validation of Patient Health Questionnaire in a Saudi sample. Ann Gen Psychiatry. 2017;16(1):32.

31. Berthold SM, Mollica RF, Silove D, Tay AK, Lavelle J, Lindert J. The HTQ-5: revision of the Harvard Trauma Questionnaire for measuring torture, trauma and DSM-5 PTSD symptoms in refugee populations. Eur J Public Health. 2019;29(3):468–74.

32. Arnetz BB, Broadbridge CL, Jamil H, Lumley MA, Pole N, Barkho E, *et al*. Specific trauma subtypes improve the predictive validity of the Harvard Trauma Questionnaire in Iraqi refugees. J Immigr Minor Health. 2014;16(6):1055–61.

33. Shoeb M, Weinstein H, Mollica R. The Harvard trauma questionnaire: adapting a cross-cultural instrument for measuring torture, trauma and posttraumatic stress disorder in Iraqi refugees. Int J Soc Psychiatry. 2007;53(5):447–63.

34. Mollica RF, Caspi-Yavin Y, Bollini P, Truong T, Tor S, Lavelle J. The Harvard Trauma Questionnaire. Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. J Nerv Ment Dis. 1992;180(2):111–6.

35. Cohen J. Kappa: coefficient of concordance. Educ Psych Meas. 1960;20(37):37–46.

36. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.

37. Mueller AE, Segal DL. Structured versus semistructured versus unstructured interviews. Encycl Clin Psychol. 2014. https://doi.org/10.1002/9781118625392.wbecp069.

38. Boardman J. Cultural adaptation of CBT for serious mental illness: a guide for training and practice by Shanaya Rathod, David Kingdon, Narsimha Pinninti, Douglas Turkington & Peter Phiri. Wiley-Blackwell. 2015.£ 32.99 (pb). 352 pp. ISBN 9781118976197. Br J Psychiatry. 2017;210(3):234–5.

39. Miranda J, Bernal G, Lau A, Kohn L, Hwang W-C, LaFromboise T. State of the science on psychosocial interventions for ethnic minorities. Annu Rev Clin Psychol. 2005;1:113–42.

40. Muñoz RF, Mendelson T. Toward evidence-based interventions for diverse populations: the San Francisco General Hospital prevention and treatment manuals. J Consult Clin Psychol. 2005;73(5):790.

41. Maercker A, Lorenz L. Adjustment disorder diagnosis: improving clinical utility. World J Biol Psychiatry. 2018;19(sup1):S3–13.

42. Zolezzi M, Alamri M, Shaar S, Rainkie D. Stigma associated with mental illness and its treatment in the Arab culture: a systematic review. Int J Soc Psychiatry. 2018;64(6):597–609.

43. Pai A, Suris AM, North CS. Posttraumatic stress disorder in the DSM-5: controversy, change, and conceptual considerations. Behav Sci. 2017;7(1):7.

44. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. Depress Anxiety. 2014;31(6):459–71.

45. Winkler JG, Brandl EJ, Bretz HJ, Heinz A, Schouler-Ocak M. Psychische Symptombelastung bei Asylsuchenden in Abhängigkeit vom Aufenthaltsstatus. Psychiatr Prax. 2019;46(04):191–9.

Publisher's Note