Published in final edited form as:

Soc Sci Med. 2017 June ; 183: 151–162. doi:10.1016/j.socscimed.2016.12.030.

How is depression experienced around the world? A systematic review of qualitative literature

E.E. Haroz^a, M. Ritchey^b, J.K. Bass^a, B.A. Kohrt^c, J. Augustinavicius^a, L. Michalopoulos^d, M.D. Burkey^e, and P. Bolton^{a,f}

^aDepartment of Mental Health, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway, Baltimore, MD 21205, United States

^bDepartment of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, 733 N Broadway, Baltimore, MD 21205, United States

^cDuke University, Duke Global Health Institute & Department of Psychiatry and Behavioral Sciences, Durham, NC 27710, United States

^dSchool of Social Work, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, United States

^eDepartment of Health, Behavior, and Society, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway, Baltimore, MD 21205, United States

^fDepartment of International Health, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St., Baltimore, MD 21205, United States

Abstract

To date global research on depression has used assessment tools based on research and clinical experience drawn from Western populations (i.e., in North American, European and Australian). There may be features of depression in non-Western populations which are not captured in current diagnostic criteria or measurement tools, as well as criteria for depression that are not relevant in other regions. We investigated this possibility through a systematic review of qualitative studies of depression worldwide. Nine online databases were searched for records that used qualitative methods to study depression. Initial searches were conducted between August 2012 and December 2012; an updated search was repeated in June of 2015 to include relevant literature published between December 30, 2012 and May 30, 2015. No date limits were set for inclusion of articles. A total of 16,130 records were identified and 138 met full inclusion criteria. Included studies were published between 1976 and 2015. These 138 studies represented data on 170 different study populations (some reported on multiple samples) and 77 different nationalities/ethnicities. Variation in results by geographical region, gender, and study context were examined to determine the consistency of descriptions across populations. Fisher's exact tests were used to compare frequencies of features across region, gender and context. Seven of the 15 features with the highest relative frequency form part of the DSM-5 diagnosis of Major Depressive Disorder (MDD). However, many of the other features with relatively high frequencies across the studies are associated features in the DSM, but are not prioritized as diagnostic criteria and therefore not included in standard instruments. The DSM-5 diagnostic criteria of *problems with concentration* and *psychomotor agitation or slowing* were infrequently mentioned. This research suggests that

the DSM model and standard instruments currently based on the DSM may not adequately reflect the experience of depression at the worldwide or regional levels.

Keywords

Depression; Systematic review; Qualitative; Cross-cultural mental health

1. Introduction

Depression is a major global public health problem. It is the leading cause of disability, with an estimated global point prevalence of 4.7% and is the eleventh leading cause of global disease burden (Ferrari et al., 2012). While the majority of the world's population lives in non-Western countries, much of the research used to describe the clinical presentation of depression has been done among populations in Western contexts (e.g., North American, European and Australian). Research and clinical work, particularly in North America, has informed the diagnostic criteria for depressive disorders in the Diagnostic and Statistical Manual, Fifth Edition (DSM-5; American Psychiatric Association, 2013) and in the International Classification of Disease, Tenth Edition (ICD-10; World Health Organization, 1992).

The DSM has its origins in attempts to track and classify mentally ill persons in institutions. The first edition of the DSM was released in 1952, largely based on psychodynamic theory, and contained 102 diagnostic categories roughly divided into conditions caused by organic brain dysfunction and conditions that resulted from the effects of the socio-cultural environment. The second edition (DSM-II) broadened the scope of the types of conditions that were included. The third edition (DSM-III) was a major turning point in American psychiatry. In response to growing criticism about what constituted true illnesses, as well as low reliability of ratings across evaluators and settings, a criteria-based classification system was developed based on the medical model of psychiatric illness. The system drew extensively from the Feighner criteria (Feighner et al., 1972) developed by research psychiatrists at Washington University in St. Louis based on the clinical experience and research of their group and others within the USA. Other criteria were agreed upon during committee meetings held by the American Psychiatric Association. Criteria for Major Depressive Disorder have remained largely unchanged since the introduction of the DSM-III. They are now nearly ubiquitous in Western psychiatric practice and research (Kawa and Giordano, 2012).

Criteria from the DSM and ICD have been widely used in psychiatric research among very different populations. In their review of 183 published studies on the mental health of refugees, Hollifield et al. (2002) found that 80% of the findings were based on instruments reflecting DSM or ICD criteria for common mental health disorders (Western instruments). Some workers in cross-cultural psychiatry and medical anthropology have criticized the accepted understanding of the construct and symptoms of depression as likely biased towards Western (i.e., North American, European and Australian) clinical populations (Bass et al., 2007; Kleinman, 2004; Summerfield, 2012). Summerfield (2012) argues that

repeatedly measuring the presence of “Western psychiatric templates [of disorders] cannot generate a universally valid knowledge base” because they may not represent the “nature of reality for the individuals under study,” (p. 5). In a meta-analysis of 20 studies, only 61% of participants endorsing culturally meaningful concepts of distress met criteria for depression (Kohrt et al., 2014), suggesting that depression criteria may miss salient features of distress cross-culturally. Taken together, this suggests that the lack of information from non-Western populations may be a significant structural barrier to a worldwide understanding of depression.

If standard instruments based on Western psychiatric nosology are biased they cannot provide an accurate picture of depression worldwide. Their use for global research and practice assumes that Western symptoms of depression are reliable and valid indicators of depression for people in other contexts (i.e. etic approach). This contrasts with an emic approach that is primarily based on local data. Some researchers have combined etic and emic approaches in their study of depression, for example to inform existing (i.e. Western) scale adaptation (e.g., Applied Mental Health Research Group (AMHR), 2013; Haroz et al., 2014; Rasmussen et al., 2014a; Rasmussen et al., 2014b). These studies began with Western-developed instruments considered to have the best face validity based on local qualitative data, then augmented with context specific symptoms using the same qualitative data. However, instruments originally developed in Western populations, even if locally adapted and tested, may still retain a substantial Western bias. There may be symptoms that are important in other parts of the world but are not emphasized or even included in the Western model, and vice versa. Systematically identifying if there are any “missed symptoms” (Summerfield, 2012, p.5) and what they are, remains an area where research is needed.

What is needed is a broader, bottom-up, open-ended approach to better understand the applicability of DSM depression diagnostic criteria and whether there are other symptoms/features that represent the common experience of depression in populations worldwide. Some researchers have done this already -- used a grounded-theory approach in which “depression” is not the starting point. Instead, they have used open-ended qualitative methods without reference to Western diagnostic criteria to better understand mental illness in the local context (Bolton, 2001; Bolton et al., 2012). However, a few isolated studies of specific populations, taken separately, are not sufficient to identify important cultural similarities and differences. A truly open ended approach covering a wide range of populations, is not feasible in a single study. Therefore, we sought to consolidate the information gathered from decades of open-ended qualitative research, in order to begin to characterize how depression is expressed and experienced in a wide variety of populations.

To the best of the authors’ knowledge there have been no systematic reviews of qualitative studies of depression features across cultures. Most previous research in this area has used quantitative data to compare differing patterns of symptom endorsement (e.g., Dere et al., 2013; Haroz et al., 2016). The reviews of qualitative studies that do exist have evaluated related topics, including perceived causes and preferred treatments for depression (Hagmayer and Engelmann, 2014), perceived barriers to accessing treatment for postpartum depression (Dennis and Chung-Lee, 2006), and perceived risk factors for postpartum depression in Sub-Saharan Africa (Wittkowski et al., 2014). Other primary research studies

have compared the expression of depression in multiple ethnic groups in a single country (c.f. Lawrence et al., 2006) and attributions and expressions of postpartum depression across sites in multiple countries (Oates et al., 2004). One of the most comprehensive cross-cultural explorations of depression is the edited book by Kleinman and Good (Kleinman, 1985) that includes chapters from clinical and anthropological perspectives examining the influences of culture in shaping depression. However, there remains no comparison of qualitative studies of depression *features* across cultural settings.

We undertook this review to address several gaps in the literature. First, this review marks an early attempt to consolidate the information on how depressive-like illness is expressed and experienced by people worldwide. Our hope is that by consolidating this information we will be able to better inform our understanding of how depression manifests across cultures. Second, we wanted to determine if concerns (e.g., potential missed symptoms/features) with cross-cultural instrument adaptation and testing were warranted, based on existing data. Findings from this review may help to improve cross-cultural measurement of depression. We intentionally used a neo-Kraepelinian approach (Compton and Guze, 1995; Rasmussen et al., 2014b) focusing solely on symptoms/features, in an effort to have our results inform self-report measurement instruments, as well as, Western psychiatric nosology as represented by the diagnostic criteria in the DSM. We intend that this review will encourage further relevant research contributing to a comprehensive picture of how depression presents and should be measured worldwide.

2. Methods

2.1. Literature search

Qualitative literature related to depression was examined through a search of peer-reviewed academic journals and solicitation of non-peer-reviewed programmatic reports related to mental health programs. The literature review followed PRISMA guidelines (Moher et al., 2009) (See Supplemental Material for PRISMA checklist). The first search was done between August 2012 and December 2012. The search strategy involved using a multi-step search of nine databases including Pubmed, Web of Science, PsycInfo, Scopus, Embase, Anthrosource, Anthropology Plus, Global Health, and Sociological Abstracts. The first step in the search used the following terms: “depression,” “depressive disorder,” “melancholia,” and “depressive disorder, major.” Once those results were returned, the second step involved reviewing this subset of records for the terms “anthropology,” “qualitative,” “ethnography,” “cross-cultural comparison,” “ethnopsychology,” “cultural characteristics,” “cross cultur*,” “phenomenology,” “cultural concepts of distress” “culture-bound” and “idioms of distress.” Study titles, abstracts, and subjects were searched and MeSH terms were used when possible. After this initial search, Google Scholar (up to the first 10,000 hits) was used to find additional references that were not identified during the initial search. Finally, the bibliography of any identified review article was searched for potentially relevant citations, and if any records not already identified were listed in the bibliography, these were included for full-text review. The process was repeated in June 2015 in order to update the search and include all records published between December 30, 2012 and May 30, 2015. No date limits were set for inclusion of articles.

Our search terms were chosen for practical reasons in an effort to efficiently identify highly relevant records. Using depression and related terms as a starting point, rather than a broader more open ended search, was done to limit the number of hits to those thought to be most relevant. However, by taking this approach, our results may be biased towards Western psychiatric nosology because we are reviewing research that has already been defined in this way.

The titles and abstracts of all records that arose during the search process were reviewed to determine if they met inclusion criteria. Authors EH and MR screened all abstracts and full texts for eligibility. Random selections of approximately 5% of abstracts and 20% of full-texts were double screened to evaluate consistency of inclusion determination. Inter-rater reliability was calculated for record eligibility determination during abstract and full-text reviews, using a Kappa statistic. A Kappa of 0 indicates less than chance agreement; 0.01–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; and 0.81–0.99 indicates almost perfect agreement (Viera and Garrett, 2005).

2.2. Inclusion criteria

Inclusion criteria were: 1) record utilized qualitative methods; 2) record had depression as the main focus; 3) record included information on features of depression; 4) record was written in English; and 5) record reported on a study population of adults between the ages of 18–65. Any records that reported data on only one individual or on a small series of case studies (i.e. less than eight people) were excluded, given the possibility that data from these studies might only reflect personal experience.

2.3. Review and data extraction

After title and abstract review, each article that met inclusion criteria was reviewed in full and the following data were extracted, if available: a) sex of the study population; b) region of the world; c) nationality and/or ethnicity; d) religious distinction; e) class distinction; f) whether the study took place in the context of peri- and post-natal, or in the context of war, trauma or displacement; g) whether the study took place in an urban or rural location; h) which qualitative research methods were used; and i) the features of depression that were mentioned in the text. The extracted features were either mentioned in direct quotes from study participants, or in summaries provided by the authors. Extracted features were classified as related to depression explicitly by authors or participants, or represented distinct descriptions of personal experience by participants in each study.

We also reviewed the methodological rigor of included studies by using criteria set out by Lincoln and Guba (1985), as well as, more specifically, the degree to which each article used a non-leading approach to data collection and the degree to which the original analysis was guided by an a priori theory. Articles were rated by four different authors independently. Each article was rated on the degree of credibility, transferability, dependability, confirmability, non-leading approach to data collection, and use of a priori theory in analysis. We used a 1- to 5-point scale for ratings; with 1 indicating very little evidence to support the criterion and 5 indicating strong evidence to support the criterion. As there are

no standard guidelines for the appraisal of the rigor of qualitative research (Thomas and Magilvy, 2011; Tobin and Begley, 2004) we did not exclude on the basis of rigor alone.

2.4. Coding

Once data were extracted from all studies, features of depression were coded based on content. We started with the diagnostic criteria for Major Depressive Disorder included in the DSM-5 (APA, 2013). Subsequent codes were added as features from each study were reviewed. Features with the same or similar meaning were grouped together. A single code representing the grouping of features was then generated. In some cases these groups contained a single symptom only (e.g., headaches). In other cases, similar concepts were grouped together and assigned a code representing the range of similar content. For example “isolation” (Bolton et al., 2012) and “social withdrawal” (Okello et al., 2012) and “lonely” (Abdur-Kadir and Bifulco, 2010) were all grouped together under the social isolation/loneliness code. Similarly, any symptom that mentioned a problem related to the heart (e.g., palpitations, weakness, heavy heart, heart pain) was grouped under “issues with the heart.” “General aches and pains” was used as a content code for any symptom that related to aches and pains beyond stomach aches and headaches. Table 1 displays examples of the codes applied and their corresponding features for the top 10 most frequent codes.

Coding of depression features was done by authors EH and JA using MaxQDA (Kuckartz, 2007) and Microsoft Excel. Any disagreements (e.g., different code used; disagreement with where features fit) within the coding process were discussed. If one of the coders was unsure about how to group a symptom, the symptom was not coded until discussed and agreement was reached about how to group it with others. For example, if study participants mentioned both being nervous and having nerves, it was discussed whether these should be lumped together or kept separate. As a reliability check, a third author (author MB) independently coded a random selection of approximately 10% of the articles. Inter-rater reliability between author MB and the other two coding authors was measured using percent agreement for each symptom code and an average percent agreement was calculated across all features codes. Kappa was not calculated for symptom codes due to the infrequency with which some codes occurred since Kappa may not be reliable for rare observations (Viera and Garrett, 2005).

After all studies had been reviewed and relevant features coded, a dataset was compiled. The dataset included rows for each study population and columns with the name of each of the content codes. As some studies reported on multiple study populations, the number of rows in the dataset was greater than the number of studies included in the review. For each study population (row), whether or not the symptom was reported was marked as present or not present (dichotomous).

2.5. Quantitative analysis

Exploratory and descriptive analyses of identified features for all study populations included in the review were performed to identify the relatively frequently mentioned features and variation by gender, study context, and geographic region. We used Fisher’s exact tests (Fisher, 1922) to examine whether frequencies of features were significantly different across

regions, by gender (male only study populations vs. female only study populations) and context (peri- and post-natal vs. all other; trauma vs. all other). Although we reported on differences in features at the $p < 0.05$ level, statistical significance was set at 0.005 after using a conservative Bonferroni correction to account for the multiple comparisons across symptom codes. Fisher's exact tests were used on a strictly exploratory basis, as sampling methods within the individual studies were not appropriate for use with inferential statistics.

3. Results

A total of 16,130 records were identified through databases searches. After initial screening, 14,960 records did not meet the inclusion criteria (i.e. were focused on children, older adults, or other disorders). One hundred ninety-seven unique records were added from the Google Scholar search, resulting in a total of 1357 records for abstract review. Fifty abstracts were double screened for eligibility by authors EH and MR. Inter-rater reliability for abstract screening was moderate ($\kappa = 0.72$; % agreement = 86%). Discrepancies in determination of article eligibility were resolved through discussion between raters.

Nine-hundred, fifty-one records were determined to be ineligible after abstract review leaving 406 articles identified for full-text review. During full-text review, 270 articles were excluded for the following reasons: not qualitative research ($n = 115$); review articles ($n = 47$); no symptom level information ($n = 43$); single case or a small case series ($n = 20$); not primarily focused on depression ($n = 25$); not published in English ($n = 10$); data reported on in another article ($n = 2$); articles could not be found through library searches ($n = 8$). In addition, two programmatic reports that had not been published in peer-reviewed literature were included in the review. Determination of inclusion from full text review was done for $n = 20$ records by authors EH and MR. Inter-rater reliability was again moderate ($\kappa = 0.79$; % agreement = 90%). Records with disagreement between raters were subsequently reviewed collaboratively and discussed to arrive at consensus about inclusion. This resulted in a total of 138 studies that met all inclusion criteria and were included in the full review (Fig. 1) (Table S1).

The authors of included studies specified a number of methods to identify features of depression and many used multiple methods ($n = 82$). According to the authors of the studies, the most common method used was in-depth and/or semi-structured interviews ($n = 95$). Other methods included focus groups ($n = 50$), key-informant interviewing ($n = 23$), ethnography ($n = 6$), the use of the Explanatory Model Interview Catalogue (EMIC) ($n = 5$), participant observation ($n = 3$), pile sort activities ($n = 3$), case vignettes ($n = 3$), illness narratives ($n = 2$), analysis of case histories ($n = 2$), photo elicitation ($n = 2$), and participatory diagramming ($n = 1$).

Studies on average used a mix of leading and non-leading approaches to data collection. The average rating for the degree to which a non-leading approach to data collection was used was 3.5 or between a mix of open-ended and leading approaches and mostly open ended approach to data collection (overall range: 1–5; average range between raters: 3.1–3.7). Only eight studies were given the lowest possible rating indicating the authors used a very leading approach to data collection. These studies represented data across regions, gender and

context and mainly focused on qualitative inquiry to help with conceptual translation of scales (e.g. Kay and Portillo, 1989; Nakimuli-Mpungu et al., 2012). Many articles were rated as being more leading because the authors purposively sampled participants based on their depressed status. For example, in a study by Okello et al. (2012), participants were identified by clinicians as having depression and then asked about their experiences. Sixty-four articles were rated as utilizing a non-leading approach to data collection.

Ratings for the degree to which a priori theory was used in analysis averaged 3.9 or mostly no a priori theory was used in analysis (overall range: 1–5; average range between raters: 3.4–4.4). Very few studies ($n = 3$) were rated with a 1, indicating a priori theory guided the analysis (i.e. use pre-specified codes). Studies that were rated with a 3, often used a mix of a priori and emergent coding. The majority of studies approached analysis without a pre-specified theory, instead utilizing techniques such as thematic analysis or content coding with emergent codes to identify features associated with depression.

Overall, it was challenging to rate articles on rigor as methodological reporting varied widely across studies. Most articles were rated as slightly above a moderate degree on credibility (average = 3.1; average range between raters: 2.9–3.8); transferability (average = 3.3; average range between raters: 2.7–3.7); and confirmability (average = 3.3; average range between raters: 3.0–3.7). The degree to which studies were dependable was lower with an average rating across studies of 2.9 (average range between raters: 2.4–3.4). Many of the articles that received low ratings provided little detail or did not report on techniques such as member-checking, reflexivity or auditing that would have enhanced the rigor of the studies.

The 138 studies represented 170 different study populations and data from 76 different nationalities/ethnicities (Fig. 2). Study populations were from North American/European/Australian non-native populations (55 study populations), Sub-Saharan Africa (38 study populations), South Asia (25 study populations), Latin America (21 study population), East Asia (7 study populations), Southeast Asia (10 study populations), the Middle East/North Africa (11 study populations), and North American/European/Australian native populations (3 study populations). No studies were identified from Russia or Central Asia. Refugee and immigrant populations were grouped together by region of origin. Sixty-five studies included data on female only samples (38.2%), eleven included data on all male only samples (6.5%), and 94 (55.3%) included data on study samples with both males and females. Twenty-seven studies focused specifically on the peri- and post-natal context (15.8%), while 25 took place in the context of war, trauma or displacement (14.7%).

Content codes based on the nine diagnostic criteria for major depression described in the DSM-5, with the addition of irritability and impaired function (i.e. eleven codes), were applied (Table 1). A total of 89 additional codes were identified during coding. There was an average of 1.5 ($SD = 2.4$; *Range*: 0–16) features that did not fit into any of the content codes (i.e. represented a unique concept) in each study population. Inter-rater reliability for the symptom content codes was calculated for $n = 15$ records. Percent agreement between raters was high (88.0%). For a full list of all codes and their frequencies see the supplemental material (Table S1).

Table 3 shows relative frequencies of depression features among all study populations ($N = 170$). *Depressed mood/sadness* ($n = 117$; 68.8%), *fatigue/loss of energy* ($n = 100$; 58.8%) and *problems with sleep* ($n = 99$; 58.2%) had the three highest relative frequencies. *Depressed mood/sadness* was described similarly across studies (see Table 2 for qualitative descriptions among select studies). Features with the highest relative frequencies across all populations ($N = 170$) were either DSM-5 diagnostic criteria or described as associated features of Major Depressive Disorder in the DSM-5. The same was true for non-Western populations only ($N = 115$). Features of depression with the highest relative frequencies that are not part of DSM-5 diagnostic criteria included *social isolation/loneliness* ($n = 92$; 54.1%), *crying a lot* ($n = 77$; 43.5%), *anger* ($n = 64$; 37.6%), *general pain* ($n = 58$; 34.1%), and *headaches* ($n = 60$; 35.3%). This was also true for non-Western populations only (i.e., excluding Northern American/European/Australian non-native populations) although the order by frequency was different (Table 3).

3.1. Results by region

As with the overall results, the features with the highest relative frequencies across regions were either DSM-5 diagnostic criteria or described as associated features of Major Depression in the DSM-5. *Depressed mood/sadness* had the highest relative frequency in Western non-indigenous, Middle Eastern/North African, and Sub-Saharan African populations ($n = 40$, 72.7%; $n = 9$, 81.8%; $n = 38$, 73.7% respectively). In Latin America and East Asia *fatigue/loss of energy* had the highest relative frequency ($n = 15$; 71.4% and $n = 6$; 85.7% respectively), while in South Asia it was *problems with sleep* ($n = 19$; 76.0%), and in Southeast Asia *depressed mood/sadness* ($n = 8$; 80.0%). Non-diagnostic features with the highest relative frequencies included: *social isolation/loneliness* in Western non-indigenous, Middle Eastern/North African and Sub-Saharan African populations ($n = 33$; 60.0%; $n = 7$, 63.6%; $n = 23$, 60.5% respectively); *crying* ($n = 13$; 61.9%) in Latin America; *general pain*, *anger*, and *anxiety* ($n = 3$; 42.9%) in East Asia; *headaches* and *issues with the heart* ($n = 13$; 52.0%) in South Asia; and *issues with the heart* ($n = 8$; 80.0%) in Southeast Asia (Table 3).

In all regions three out of the top five most frequently mentioned features are DSM-5 symptom criteria of MDD (Table 4). Associated (non-diagnostic) DSM-5 features had the highest relative frequency only in Southeast Asian populations. Statistically significant differences ($p < 0.005$) in relative frequencies across regions were observed for the following features: *headaches*, *worry*, *thinking too much*, *heart issues*, *anxiety*, *weakness*, *dizziness*, *feelings of blackness*, *trouble breathing*, *disappointed*, *pressure in the chest* and *digestion* (Table S2). Other features such as, *low self-esteem*, *feeling scared*, *confusion*, as well as others, were significant at the $p < 0.05$ level, but were not statistically significantly different across regions after adjusting for multiple comparisons (Table S2 and S3).

3.2. Ubiquitous features

Fig. 3 is a Venn-like diagram, with the rings representing each study region (with the exception of the Western indigenous populations because of the small sample size) and the center circle representing features that arose during the review and present in at least one study population from these regions. Features in the table were ranked by their relative

frequency by taking the product of their frequencies in each region. Thirty-six features (out of 100 features) appeared in every region. All eleven DSM-5 diagnostic features were present in every region, while most of the other 25 were included in the DSM-5 as associated features.

3.3. Results by gender

Among female-only study populations ($n = 65$) four out of the five features with the highest relative frequencies were DSM-5 diagnostic criteria for MDD and included *depressed mood/sadness* ($n = 42$; 64.6%), *fatigue/loss of energy* ($n = 39$; 60.0%), *problems with sleep* ($n = 37$; 56.9%), and *weight/appetite issues* ($n = 25$; 38.5%). *Social isolation/loneliness* ($n = 33$; 50.8%), *general pain* ($n = 23$; 35.4%), *headaches* ($n = 22$; 33.8%), and *crying* ($n = 23$; 35.4%) also had relatively high frequencies. The number of studies involving all-male populations was small ($n = 11$). *Depressed mood/sadness* ($n = 8$; 72.7%), *anger* ($n = 6$; 54.5%) *social isolation/loneliness* ($n = 6$; 54.5%), and *weight/appetite problems* ($n = 6$; 54.5%) had the highest relative frequencies (Table 5). When comparing frequencies of features across genders, no statistically significant differences ($p < 0.005$) were detected. However, at the $p < 0.05$ level, several features emerged as trending towards significantly different between men and women including: *substance use/abuse*, *staying in bed*, and *aggression* which were all higher in all male populations compared to all female populations (Table S3). Despite the small number of male-focused studies this review suggests that subjective experiences of depression are similar across genders.

3.4. Results by context

Results by contextual variable are presented in Table 6. In the context of trauma, most study populations reported *problems with sleep* ($n = 17$; 68.0%), *social isolation/loneliness* ($n = 16$; 64.0%), *depressed mood/sadness* ($n = 15$; 60.00%), and *weight/appetite problems* ($n = 15$; 60.0%). In peri- and post-natal contexts, *social isolation/loneliness* ($n = 19$; 70.4%) was the most common symptom, followed by *depressed mood/sadness* ($n = 18$; 66.6%), and *fatigue/loss of energy* ($n = 15$; 55.5%).

The frequency of features was not significantly ($p < 0.005$) different between study populations when comparing peri- and post-natal populations to all others. However, some features appeared to be trending ($p < 0.05$) towards significance, including *general pain*, *not talking to others*, *substance use/abuse* and *homicidal thoughts*. These results suggest that features from peri- and post-natal populations are relatively similar to overall populations.

Among trauma-affected populations, the frequencies of *hopelessness*, *rumination*, and *feeling suspicious* were significantly different ($p < 0.005$) than the frequencies observed in other populations. A greater proportion of study populations in trauma-affected populations reported feeling these features compared to all other populations. Other features that were trending towards significantly different ($p < 0.05$) included *problems with sleep*, *weight/appetite problems*, *crying*, *impaired functioning*, *not talking to others*, *problems with memory*, *feeling trapped*, *lack of coping*, *feeling regretful*, *nightmares*, and *feelings of grief*. Many of these features overlap with posttraumatic stress and may be particularly salient given this population's higher risk for this disorder.

4. Discussion

Most research on depression conducted among non-Western populations has used measurement instruments and diagnostic criteria based on the DSM and ICD presentations of depression. This assumes that these criteria, developed among Western populations, are applicable across cultures. We reviewed available worldwide qualitative studies on depression in order to explore the evidence for this assumption, as well as to provide a less culturally biased understanding of how depression is experienced across populations.

The DSM-5 diagnostic criteria for MDD were reported across all regions, genders and socio-cultural contexts. Most other frequently described features also appear in the DSM-5 as associated features of MDD. Across all regions, features of *depressed mood/sadness*, *fatigue/loss of energy*, *problems with sleep*, *appetite/weight problems*, *suicidal thoughts*, *loss of interest*, and *worthlessness/guilt* were commonly reported in qualitative studies of depression, with *irritability* also frequent to a lesser degree than the other features. Most of the remaining frequently described features also appear in the DSM-5 as associated features of MDD.

Our findings do, however, suggest problems with the cross-cultural use of the DSM-5 Major Depression diagnostic criteria. Four of the most frequently mentioned ubiquitous features across studies were not part of DSM-5 diagnostic criteria: *social isolation/loneliness*, *crying*, *anger*, and *general pain*. In contrast, other DSM-5 diagnostic features were not frequently reported in the global literature, specifically, *problems with concentration* and *psychomotor agitation or slowing*.

Many features overlapped with symptoms of anxiety disorders, such as *worry*, *issues with breathing*, *irritability*, *problems with sleep*, and *restlessness*. The DSM-5 currently separates Major Depressive Disorder and Generalized Anxiety Disorder. However, there is a robust evidence showing that depression and anxiety are often co-morbid (Kessler et al., 2008), share similar risk factors (Almeida et al., 2012), exhibit similar neurocognitive processes involving the limbic system (Ressler and Nemeroff, 2000), and respond to similar treatments (Butler et al., 2006). While we limited our search strategy to records with a main focus of depression, we did not exclude records that included anxiety, as long as depression was one of the main foci of the record. Results from our review support previous literature which suggests that the symptoms of depression and anxiety demonstrate substantial overlap and are highly comorbid (Abas and Broadhead, 1997; Bener et al., 2012; Das-Munshi et al., 2008; Kaaya et al., 2002).

Anger was frequently mentioned overall ($n = 61$; 36.3%), and was present with comparable frequency in both female-only populations ($n = 16$; 25.8%) and male-only populations ($n = 3$; 37.5%). There is general consensus that depression in men usually consists of symptoms of anger, impoverished social relationships, emotional numbness, impulse control difficulties, irritability, aggression, substance use, and suicide (Brownhill et al., 2005; Cochran and Rabinowitz, 2003; Martin et al., 2013; Oliffe and Phillips, 2008). However, results from the current review indicate that women around the world also commonly report *anger* as a symptom of depression, which has been found in other studies (Rees et al., 2013;

Williamson et al., 2014). Anger may represent a common sign of depression in women that has not been fully acknowledged in psychiatric classification.

Somatic complaints were very common among all study populations, including Western populations. Reviews of other research suggest that somatic complaints related to depression are ubiquitous worldwide (Draguns and Tanaka-Matsumi, 2003; Shidhaye et al., 2013) and their expression often functions as a reflection both of the individual and of problems with the broader healthcare system (Kirmayer, 2001). People may express these complaints as a method for getting help with their distress when the overall system does not understand or provide services for more cognitive or behavioral symptoms, as suggested by the literature on idioms of distress which are commonly somatic (Nichter, 2010).

Some features appear to be contextually specific and may warrant inclusion in measurement instruments used among specific populations. For example, in trauma-affected study populations, features that overlap with PTSD, such as guilt, rumination, and feeling suspicious occur more frequently compared to all populations combined. Consistent with other research, there is high overlap with depression and PTSD symptomology in populations affected by collective trauma (Momartin et al., 2004; O'Donnell et al., 2004). Other features that arose more frequently in some populations such as *worry* (more common in South Asian and Southeast Asian populations) or *thinking too much* (more common in Southeast Asian and Sub-Saharan African populations), may be important to include in measurement instruments for use in these particular regions.

4.1. Limitations

Our review was limited to studies published in the English language literature so we have likely missed relevant studies. Features extracted in this review were also already translated into English by the authors of the studies and it is possible that during translation nuances of the literal expressions of the features were not captured. Limiting records to those in English may have biased the results in favor of DSM features since English speakers seeking to better understand depression in other cultures may be primed to look specifically for people who fit their conceptualization of depression.

While our goal was to avoid this bias by selecting open-ended, qualitative literature the degree to which studies achieved this varied. All studies utilized qualitative methods, and most involved data collection and analysis using open-ended approaches. However, there were a small number of studies included that used qualitative methods to confirm a priori assumptions. Moreover, assessing the rigor of methods was challenging as the adequacy of methodological reporting varied significantly across studies. Many studies did not report clearly on aspects of credibility, transferability, dependability, or confirmability, which reduced our ability to assess their rigor and that of the studies overall.

We did not include grey literature in our search strategy but did solicit reports of studies from researchers working in the field that had not been published in peer-reviewed literature. Only two records from the non-peer reviewed grey literature were found and included. Our search strategy did not explicitly include “mixed-methods” perhaps resulting in missed features as well. While a proportion of the included studies represented mixed-methods

studies, future research should explicitly examine this body of literature more thoroughly. Another limitation is in our exclusion of case studies. Some case studies may reflect a rich description of depression that would not have otherwise been captured in broader qualitative work. However, since case studies include data on such a limited number of participants, that they may not represent the experiences of the more general population. The results from the Fisher exact tests are suggestive only, due to the likelihood of sampling bias within the individual studies.

There is a potential tautological error that the search for depression features across global populations will result with confirmation of the diagnosis because most studies are influenced by psychiatric categorizations (Kleinman, 1988). Thus, DSM/ICD features will typically be more thoroughly evaluated through qualitative interview prompts and more likely to be reported than features not associated with the psychiatric diagnosis of depression. Limiting our review to studies explicitly focused on depression increases this bias. Search terms were selected for practical reasons in order to limit the number of potential hits to articles that would be relevant to our research question. We acknowledge this bias and the subsequent over-emphasis on existing diagnostic criteria in the results extracted. Given this bias, we do find it striking that some features, though nearly ubiquitous across populations, are not part of current DSM diagnostic criteria: *loneliness, anger, crying*, and somatic complaints. Their absence from instruments based on the DSM criteria may constitute a gap in the assessment of depression symptomatology generally while the absence of regionally or locally important features may constitute a bias in cross-national studies.

A second conceptual limitation is that the current study does not address the personal and cultural significance of the features we have extracted. This could lead to “category fallacies”, i.e., assumptions that a symptom or group of symptoms will have comparable personal and cultural meaning across social groups (Kleinman, 1977). Potential mislabeling of features associated with disorder in one culture as also being a ‘disorder’ in another culture (in the absence of evidence of prolonged impairment) is a major concern in anthropology and cross-cultural psychiatry in terms of the potential for pathologizing normative non-disabling behavior, i.e., medicalization (de Jong and Reis, 2013; Kleinman, 2008; Nichter, 2010).

One way to reduce the risk of this category fallacy is to examine if features are associated with expected life impact. In this study, we evaluated the relationship between features and impaired functioning as an indicator of comparable impact across cultural groups. Unfortunately, problems with daily functioning were only mentioned explicitly in one quarter of the study populations included in this review. The majority of study populations did not raise problems with daily functioning as part of their subjective experiences of depression. To the extent possible, future studies and literature synthesis should not only focus on the presence of features but also the meaning of these experiences and the association with functional impairment (c.f., Bolton and Tang, 2002). This focus would help to distinguish between depression, conceived as a psychiatric disorder, and culturally-normative displays of sadness and grief that do not lead to prolonged impairment (Horwitz

and Wakefield, 2007). Use of specific probes during data collection or more reporting on findings related to impairment would help to make this distinction.

Multiple approaches can be used to understand cross-cultural and cross-population commonalities in expression of distress. A complimentary strategy to our symptom-based approach is to examine cultural concepts of distress, that are shared across groups. A recent review on cultural concepts of distress related to “thinking too much” identified 138 publications on this manifestation of distress across more than 20 countries covering Asia, Africa, Europe, North and South America, and Australia/Pacific Islands (Kaiser et al., 2015). Thinking too much was associated with low mood, anhedonia, poor concentration, social withdrawal, sleep disruptions, and somatic complaints across most populations. Moreover, manifestations of thinking too much were related to functional impairment and often perceived to result in more severe mental health problems if not formally or informally treated. This approach started with locally salient categories then explored potential overlap with psychiatric categories such as depression, anxiety, and PTSD. Ultimately, both approaches (i.e., reviewing both the literature on depression and cultural concepts of distress) are necessary and complimentary to fully understand commonalities and diversity in experiencing psychological distress.

5. Conclusions

We undertook this review to determine whether the current practice of using standard depression instruments based on Western models (specifically the DSM) may represent a significant bias in global mental health research, even after local adaptation and testing. We found that the diagnostic and associated features described in the DSM-5 are consistent with the frequently mentioned features in our review at both the regional and worldwide levels. However, worldwide and at the regional level, some DSM non-diagnostic features were mentioned by more study populations than were some diagnostic features, while several diagnostic features were not prominent in some regions. These findings suggest a need for review of the content of standard instruments beyond their current focus on DSM diagnostic criteria, in order to accurately reflect the experience of depression worldwide and particularly for non-Western populations. Our findings also support the need for regional variation in instruments in accord with local variation in presentation.

While we have confidence in our approach, the robustness of our findings is limited by the small number of studies we could find and the limited information available on how they were conducted. We propose to repeat this process as more qualitative studies become available. We therefore advocate for a coordinated effort to conduct relevant and high quality qualitative research on how depression manifests across cultures. Meanwhile, this paper describes the first attempt to review the available qualitative literature to better understand how depression manifests worldwide, in order to develop more appropriate instruments for both worldwide and regional use. Based on this review, we believe instruments that only include DSM-5 diagnostic criteria are inadequate for use in populations world-wide. These, instruments should not only include features representing diagnostic criteria but also incorporate highly relevant associated features such as *social isolation or loneliness*, *excessive crying*, *anger*, and *general aches and pain*. Patterns of regional variation

documented here can help guide researchers and clinicians to features relevant for specific populations and settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

United States Agency for International Development/Victims of Torture Fund AID-DFD-A-00-08-00308.

National Institute of Mental Health: T32 MH014592-38. NIMH grant number: K01MH104310

The authors would like to thank Dr. Wietse Tol for help with development of the search strategy.

References

- Abas MA, Broadhead JC. Depression and anxiety among women in an urban setting in Zimbabwe. *Psychol Med*. 1997; 27(01):59–71. [PubMed: 9122309]
- Abdur-Kadir, Bifulco A. Malaysian muslim mothers' experience of depression and service use. *Cult Med Psychiatry*. 2010; 34(3):443–467. [PubMed: 20549550]
- Almeida OP, Draper B, Pirkis J, Snowdon J, Lautenschlager NT, Byrne G, Flicker L. Anxiety, depression, and comorbid anxiety and depression: risk factors and outcome over two years. *Int Psychogeriatr*. 2012; 24(10):1622–1632. [PubMed: 22687290]
- Amankwaa LC. Postpartum depression, culture and african-american women. *J Cult Divers*. 2003; 10(1):23–29. [PubMed: 12776544]
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders Fifth Edition*. 5. Author; Washington, DC: 2013.
- Applied Mental Health Research Group (AMHR). *Design, Implementation, Monitoring and Evaluation of Cross-cultural Trauma Related Mental Health and Psychosocial Assistance Programs: a User's Manual for Researchers and Program Implementers*. 2013 Retrieved from.
- Bass JK, Bolton PA, Murray LK. Do not forget culture when studying mental health. *Lancet*. 2007; 370(9591):918–919. [PubMed: 17869621]
- Bener A, Ghuloum S, Abou-Saleh MT. Prevalence, symptom patterns and comorbidity of anxiety and depressive disorders in primary care in Qatar. *Soc Psychiatry Psychiatric Epidemiol*. 2012; 47(3): 439–446.
- Bernstein KS, Lee J, Park S, Jyoung J. Symptom manifestations and expressions among korean immigrant women suffering with depression. *J Adv Nurs*. 2008; 61(4):393–402. [PubMed: 18234037]
- Bolton P. Local perceptions of the mental health effects of the rwandan genocide. *J Nerv Ment Dis*. 2001; 189(4):243–248. [PubMed: 11339320]
- Bolton P, Michalopoulos L, Ahmed AM, Murray LK, Bass J. The mental health and psychosocial problems of survivors of torture and genocide in Kurdistan, Northern Iraq: a brief qualitative study. *Torture Q J Rehabil Torture Vict Prev Torture*. 2012; 23(1):1–14.
- Bolton P, Tang AM. An alternative approach to cross-cultural function assessment. *Soc Psychiatry Psychiatric Epidemiol*. 2002; 37(11):537–543.
- Borra R. Depressive disorder among Turkish women in the Netherlands: A qualitative study of idioms of distress. *Transcult Psychiatr*. 2011; 48(5):660–674.
- Brownhill S, Wilhelm K, Barclay L, Schmied V. 'Big build': hidden depression in men. *Aust N Z J Psychiatry*. 2005; 39(10):921–931. [PubMed: 16168020]
- Brownhill S, Wilhelm K, Barclay L, Parker G. Detecting depression in men: A matter of guesswork. *Int J Mens Health*. 2002; 1(3):259–271.

- Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clin Psychol Rev*. 2006; 26(1):17–31. [PubMed: 16199119]
- Cochran SV, Rabinowitz FE. Gender-sensitive recommendations for assessment and treatment of depression in men. *Prof Psychol Res Pract*. 2003; 34(2):132–140.
- Compton WM, Guze SB. The neo-Kraepelinian revolution in psychiatric diagnosis. *Eur Archives Psychiatry Clin Neurosci*. 1995; 245(4–5):196–201.
- Das-Munshi J, Goldberg D, Bebbington PE, Bhugra DK, Brugha TS, Dewey ME, Prince M. Public health significance of mixed anxiety and depression: beyond current classification. *Br J Psychiatry*. 2008; 192(3):171–177. [PubMed: 18310574]
- Danielsson U, Johansson EE. Beyond weeping and crying: a gender analysis of expressions of depression. *Scand J Prim Health Care*. 2005; 23(3):171–177. [PubMed: 16162470]
- de Jong J, Reis R. Collective trauma processing: dissociation as a way of processing postwar traumatic stress in Guinea Bissau. *Transcult Psychiatry*. 2013; 50(5):644–661. 1363461513500517. [PubMed: 23999989]
- Dejman M, Forouzan AS, Assari S, Malekafzali H, Nohesara S, Khatibzadeh N, et al. An explanatory model of depression among female patients in Fars, Kurds, Turks ethnic groups of Iran. *Iranian J Public Health*. 2011; 40(3):79–88.
- Dennis CL, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. *Birth*. 2006; 33(4):323–331. [PubMed: 17150072]
- Dere J, Sun J, Zhao Y, Persson TJ, Zhu X, Yao S, Ryder AG. Beyond “somatization” and “psychologization”: symptom-level variation in depressed Han Chinese and Euro-Canadian outpatients. *Front Psychol*. 2013; 4
- Draguns J, Tanaka-Matsumi J. Assessment of psychopathology across and within cultures: issues and findings. *Behav Res Ther*. 2003; 41(7):755–776. [PubMed: 12781244]
- Feighner JP, Robins E, Guze SB, Woodruff RA, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. *Archives General Psychiatry*. 1972; 26(1):57–63.
- Ferrari A, Somerville A, Baxter A, Norman R, Patten S, Vos T, Whiteford H. Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature. *Psychol Med*. 2012:1–11.
- Fisher RA. On the interpretation of χ^2 from contingency tables, and the calculation of P. *J R Stat Soc*. 1922; 85(1):87–94.
- Halbreich U, Alarcon RD, Calil H, Douki S, Gaszner P, Jadresic E, Trivedi JK. Culturally-sensitive complaints of depressions and anxieties in women. *J Affect Disord*. 2007; 102(1–3):159–176. [PubMed: 17092564]
- Hagmayer Y, Engelmann N. Causal beliefs about depression in different cultural groups—what do cognitive psychological theories of causal learning and reasoning predict? *Front Psychol*. 2014; 5
- Haroz EE, Bolton P, Gross A, Chan KS, Michalopoulos L, Bass J. Depression symptoms across cultures: an IRT analysis of standard depression symptoms using data from eight countries. *Soc Psychiatry Psychiatric Epidemiol*. 2016:1–11.
- Haroz EE, Bass JK, Lee C, Murray LK, Robinson C, Bolton P. Adaptation and testing of psychosocial assessment instruments for cross-cultural use: an example from the Thailand Burma border. *BMC Psychol*. 2014; 2(1):31–40. [PubMed: 25685351]
- Hollifield M, Warner TD, Lian N, Krakow B, Jenkins JH, Kesler J, Westermeyer J. Measuring trauma and health status in refugees: a critical review. *JAMA*. 2002; 288(5):611–621. [PubMed: 12150673]
- Horwitz, AV., Wakefield, JC. *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford University Press; 2007.
- James S, Navara GS, Clarke JN, Lomotey J. An inquiry into the “agonies” (agonias) of Portuguese immigrants from the Azores. *Hisp J Behav Sci*. 2005; 27(4):547–564.
- Jayawickreme N, Jayawickreme E, Goonasekera MA, Foa EB. Distress, wellbeing and war: Qualitative analyses of civilian interviews from north eastern Sri Lanka. *Intervention*. 2009; 7(3): 204–222.

- Kaaya SF, Fawzi M, Mbwapbo J, Lee B, Msamanga GI, Fawzi W. Validity of the hopkins symptom Checklist-25 amongst HIV-positive pregnant women in Tanzania. *Acta Psychiatr Scand*. 2002; 106(1):9–19. [PubMed: 12100343]
- Kaaya SF, Mbwapbo JK, Smith Fawzi MC, Van DB, Schaalma H, Leshabari MT. Understanding women's experiences of distress during pregnancy in Dar es Salaam, Tanzania. *Tanzan J Health Res*. 2010; 12(1):36–46. [PubMed: 20737827]
- Kaiser BN, Haroz EE, Kohrt BA, Bolton PA, Bass JK, Hinton DE. "Thinking too much": a systematic review of a common idiom of distress. *Soc Sci Med*. 2015; 147:170–183. [PubMed: 26584235]
- Kaiser BN, McLean KE, Kohrt BA, Hagaman AK, Wagenaar BH, Khoury NM, et al. *Reflechi twòp—Thinking too much: Description of a cultural syndrome in haiti's central plateau*. *Cult Med Psychiatr*. 2014; 38(3):448–472.
- Kay M, Portillo C. Nervios and dysphoria in Mexican American widows. *Health Care Women Int*. 1989; 10(2–3):273–293. [PubMed: 2768095]
- Kawa S, Giordano J. A brief historicity of the Diagnostic and Statistical Manual of Mental Disorders: issues and implications for the future of psychiatric canon and practice. *Philos Ethics Humanit Med*. 2012; 7(1):1. [PubMed: 22243957]
- Kemp M. Hearts and minds: Agency and discourse on distress. *Anthropol Med*. 2003; 10(2):187–205. [PubMed: 26954836]
- Kessler RC, Gruber M, Hettema JM, Hwang I, Sampson N, Yonkers KA. Co-morbid major depression and generalized anxiety disorders in the national comorbidity survey follow-up. *Psychol Med*. 2008; 38(03):365–374. [PubMed: 18047766]
- Kirmayer L. Cultural/variations in the clinical presentation of depression and anxiety: implications for diagnosis and treatment. *J Clin Psychiatry*. 2001; 62(13):22–30.
- Kleinman, A. *Culture and Depression: Studies in the Anthropology and Cross-cultural Psychiatry of Affect and Disorder*. Vol. 16. Univ of California Press; 1985.
- Kleinman A. Culture and depression. *N Engl J Med*. 2004; 351(10):951–953. [PubMed: 15342799]
- Kleinman, A. *Rethinking Psychiatry*. Simon and Schuster; 2008.
- Kleinman, A. *Rethinking Psychiatry: from Cultural Category to Personal Experience*. Free Press; Collier Macmillan; New York: 1988.
- Kleinman AM. Depression, somatization and the "new cross-cultural psychiatry". *Soc Sci Med*. 1967; 11(1):3–9. 1977.
- Kohrt BA, Rasmussen A, Kaiser BN, Haroz EE, Maharjan SM, Mutamba BB, Hinton DE. Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. *Int J Epidemiol*. 2014; 43(2):365–406. [PubMed: 24366490]
- Koo K. Carers' representations of affective mental disorders in British Chinese communities. *Sociol Health Illn*. 2012; 34(8):1140–1155. [PubMed: 22332911]
- Kuckartz, U. *MAXQDA: Qualitative Data Analysis*. VERBI software; Berlin: 2007.
- Lawrence V, Murray J, Banerjee S, Turner S, Sangha K, Byng R, Macdonald A. Concepts and causation of depression: a cross-cultural study of the beliefs of older adults. *Gerontologist*. 2006; 46(1):23–32. [PubMed: 16452281]
- Lazear KJ, Pires SA, Isaacs MR, Chaulk P, Huang L. Depression among low-income women of color: Qualitative findings from cross-cultural focus groups. *J Immigr Minor Health*. 2008; 10(2):127–133. [PubMed: 18236157]
- Lim AG, Stock L, Shwe Oo EK, Jutte DP. Trauma and mental health of medics in eastern Myanmar's conflict zones: A cross-sectional and mixed methods investigation. *Confl Health*. 2013; 7(1):15–28. [PubMed: 23899166]
- Lincoln, YS., Guba, EG. *Naturalistic Inquiry*. Sage Publications; Newbury Park, CA: 1985.
- Mallinson S, Popay J. Describing depression: Ethnicity and the use of somatic imagery in accounts of mental distress. *Sociol Health Illn*. 2007; 29(6):857–871. [PubMed: 17986019]
- Martin LA, Neighbors HW, Griffith DM. The experience of symptoms of depression in men vs women: analysis of the national comorbidity survey replication. *JAMA Psychiatry*. 2013; 70(10):1100–1106. [PubMed: 23986338]

- Martinez Tyson DD, Castañeda H, Porter M, Quiroz M, Carrion I. More similar than different? Exploring cultural models of depression among Latino immigrants in Florida. *Depress Res Treat*. 2011;1–11.
- Meffert SM, Marmar CR. Darfur refugees in Cairo mental health and interpersonal conflict in the aftermath of genocide. *J Interpers Violence*. 2009; 24(11):1835–1848. [PubMed: 18945917]
- Meyer SR, Robinson WC, Chhim S, Bass JK. Labor migration and mental health in Cambodia: a qualitative study. *J Nerv Ment Dis*. 2014; 202(3):200–208. [PubMed: 24566505]
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009; 151(4):264–269. [PubMed: 19622511]
- Momartin S, Silove D, Manicavasagar V, Steel Z. Comorbidity of PTSD and depression: associations with trauma exposure, symptom severity and functional impairment in Bosnian refugees resettled in Australia. *J Affect Disord*. 2004; 80(2):231–238. [PubMed: 15207936]
- Muhwezi WW, Okello ES, Neema S, Musisi S. Caregivers' experiences with major depression concealed by physical illness in patients recruited from central Ugandan Primary Health Care Centers. *Qual Health Res*. 2008; 18(8):1096–1114. [PubMed: 18650565]
- Murray L, Bass J, Bolton P. Qualitative Study to Identify Indicators of Psychological Problems and Functional Impairment Among Residents of Sange District. South Kivu, Eastern DRC, A report to the Victims of Torture Fund United States Agency for International Development (USAID). 2006
- Nakimuli-Mpungu E, Mojtabai R, Alexandre PK, Katabira E, Musisi S, Nachega JB, Bass JK. Cross-cultural adaptation and validation of the self-reporting questionnaire among HIV individuals in a rural ART program in southern Uganda. *HIV AIDS Auckl*. 2012; 4:51–60. [PubMed: 22570575]
- Naem F, Ayub M, Kingdon D, Gobbi M. Views of depressed patients in Pakistan concerning their illness, its causes, and treatments. *Qual Health Res*. 2012; 22(8):1083–1093. [PubMed: 22707343]
- Nichter M. Idioms of distress revisited. *Cult Med Psychiatry*. 2010; 34(2):401–416. [PubMed: 20495999]
- Newsma JA. Indigenous perspectives on depression in rural regions of India and the United States. *Transcult Psychiatr*. 2011; 48(5):539–568.
- O'Donnell ML, Creamer M, Pattison P. Posttraumatic stress disorder and depression following trauma: understanding comorbidity. *Am J Psychiatry*. 2004; 161(8):1390–1396. [PubMed: 15285964]
- Oates MR, Cox JL, Neema S, Asten P, Glangeaud-Freudenthal N, Figueiredo B, Kammerer MH. Postnatal depression across countries and cultures: a qualitative study. *Br J Psychiatry*. 2004; 184(46):s10–s16.
- Okello ES, Ngo VK, Ryan G, Musisi S, Akena D, Nakasujja N, Wagner G. Qualitative study of the influence of antidepressants on the psychological health of patients on antiretroviral therapy in Uganda. *Afr J AIDS Res*. 2012; 11(1):27–44. [PubMed: 25870895]
- Oliffe JL, Phillips MJ. Men, depression and masculinities: a review and recommendations. *J Men's Health*. 2008; 5(3):194–202.
- Poudyal B, Bass J, Subyantoro T, Jonathan A, Erni T, Bolton P. Assessment of the psychosocial and mental health needs, dysfunction and coping mechanisms of violence affected populations in Bireuen, Aceh. *Torture: Q J Rehabil Torture Vict Prev Torture*. 2009; 19(3):218–226.
- Rasmussen A, Eustache E, Raviola G, Kaiser B, Grelotti DJ, Belkin GS. Development and validation of a Haitian Creole screening instrument for depression. *Transcult Psychiatry*. 2014; 52(1):33–57. [PubMed: 25080426]
- Rasmussen A, Keatley E, Joscelyne A. Posttraumatic stress in emergency settings outside North America and Europe: a review of the emic literature. *Soc Sci Med*. 2014; 109:44–54. [PubMed: 24698712]
- Rees S, Silove D, Verdial T, Tam N, Savio E, Fonseca Z, Tay K. Intermittent explosive disorder amongst women in conflict affected Timor-Leste: associations with human rights trauma, ongoing violence, poverty, and injustice. *PloS One*. 2013; 8(8):e69207. [PubMed: 23950885]
- Rees S, Silove D. Sakit Hati: A state of chronic mental distress related to resentment and anger amongst West Papuan refugees exposed to persecution. *Soc Sci Med*. 2011; 73(1):103–110. [PubMed: 21665342]

- Ressler KJ, Nemeroff CB. Role of serotonergic and noradrenergic systems in the pathophysiology of depression and anxiety disorders. *Depress Anxiety*. 2000; 12(S1):2–19. [PubMed: 11098410]
- Shidhaye R, Mendenhall E, Sumathipala K, Sumathipala A, Patel V. Association of somatoform disorders with anxiety and depression in women in low and middle income countries: a systematic review. *Int Rev Psychiatry*. 2013; 25(1):65–76. [PubMed: 23383668]
- Sulaiman SOY, Bhugra D, De Silva P. The development of a culturally sensitive symptom checklist for depression in Dubai. *Transcult Psychiatr*. 2001; 38(2):219–229.
- Summerfield D. Afterword: against global mental health. *Transcult Psychiatry*. 2012; 49(3):519. [PubMed: 23008353]
- Thomas E, Magilvy JK. Qualitative rigor or research validity in qualitative research. *J Spec Pediatr Nurs*. 2011; 16(2):151–155. [PubMed: 21439005]
- Tobin GA, Begley CM. Methodological rigour within a qualitative framework. *J Adv Nurs*. 2004; 48(4):388–396. [PubMed: 15500533]
- Ventevogel P, Jordans M, Reis R, de Jong J. Madness or sadness? Local concepts of mental illness in four conflict-affected African communities. *Confl Health*. 2013; 7(1):3–19. [PubMed: 23418727]
- Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005; 37(5):360–363. [PubMed: 15883903]
- Williamson JA, O'Hara MW, Stuart S, Hart KJ, Watson D. Assessment of postpartum depressive symptoms: the importance of somatic symptoms and irritability. *Assessment*. 2014; 22(3):309–318. [PubMed: 25059681]
- Wittkowski A, Gardner PL, Bunton P, Edge D. Culturally determined risk factors for postnatal depression in Sub-Saharan Africa: a mixed method systematic review. *J Affect Disord*. 2014; 163:115–124. [PubMed: 24461216]
- World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. World Health Organization; Geneva: 1992.

Appendix A. Supplementary data

The following is the supplementary data related to this article:

Research Highlights

1. Reviews qualitative literature to identify features of depression worldwide.
2. Emphasizes importance of qualitative data to enhance understanding of depression.
3. Investigates potential biases in Western instruments that measure depression.
4. Argues that Western diagnostic criteria do not represent depression globally.

Figure 1.

N = 170 study populations representing 76 different nationalities and ethnicities

* Dots represent nationalities/ethnicities

Figure 2.
Geographic variation of study populations*.

*Bold indicates included in DSM-5 diagnostic criteria for Major Depression

Figure 3.
Ubiquitous features associated with depression.

Table 1

Most frequently applied codes and examples of corresponding features.

Content code	Author, Year	Text from articles ^a
Depressed Mood/Sadness	Dejman et al., 2011	“afsordehgi,” (depression) “darikma” (deep sadness with anxiety)
	Lim et al., 2013	“Seik da’kya” (depression or spirit fall)
	Meffert and Marmar, 2009	“depressed,” “very sad”
Fatigue/loss of energy	Kaaya et al., 2010	“loss of energy
	Martinez Tyson, 2011	“Tiredness,” “fatigue”
	Naeem et al., 2012	“fatigue”
Problems with sleep	Kemp, 2003	“don’t sleep well,” “I can be up for the whole night”
	Muhwesi et al., 2008	“lack of sleep”
	Murray et al., 2006	“losing sleep”
Social isolation/loneliness	Abdur-Kadir and Bifulco, 2010	“alone” “lonely”
	Danielsson and Johansson, 2005	“feelings of loneliness”
	Ventevogel et al., 2013	“social isolation”
Appetite/weight problems	Bernstein et al., 2008	“losing a taste for food, “lost appetite, “can’t eat”
	James et al., 2005	“I could not eat”
	Lazear et al., 2008	“changes in appetite and weight”
Crying a lot	Halbreich et al., 2007	“on the verge of crying” (India) “got crying spells for nothing” (Brazil)
	Okello et al., 2012	“crying”
	Rees and Silove, 2011	“crying”
Loss of interest	Jayawickreme et al., 2009	“less interest in things outside the home”
	Koo, 2012	“nothing interests him anymore”
	Nieuwsma, 2011	“lack of interest in everything”
Suicidal thoughts	Amankwaa, 2003	“contemplating suicide”
	Bolton et al., 2012	“feeling you would be better dead”
	Brownhill et al., 2002	“having suicidal thoughts or attempting suicide”
General aches and pains	Borra, 2011	“neck and shoulder pains,” “back pains”
	Okello et al., 2012	“pain in all the joints”
	Sulaiman et al., 2001	“aches and pains” (men)
Anger	Mallinson and Popay, 2007	“angry”
	Poudyal et al., 2009	“easily get angry”
	Ventevogel et al., 2013	“easily angry”

^aWhere possible included text in original language.

Table 2

Features commonly associated with depressed mood/sadness from select studies.

Study	Associated features of depressed mood/sadness
Bolton et al., 2012	“thinks a lot with hand on chin,” “makes people sick/take to their bed,” “person is unstable/unbalanced/cannot think straight,” “cannot sleep,” “food does not taste good/unable to eat,” “because of sadness the person stays alone”
Halbreich et al., 2007	loss of pleasure/interest in their job, risk-taking behaviors, irritability, thoughts of suicide, helplessness, difficulty concentrating, fatigue, loss of desire for pleasure
Kaiser et al., 2014	“thinking too much,” “distance in one’s thoughts,” “becoming quiet, cannot cry,” “sadness really lies in one’s thoughts”
Lim et al., 2013	“spirit fall,” “the heart falls down”
Meyer et al., 2014	“can’t sleep or eat,” “thinking too much,” “afraid of the boss,” “not earning enough money,” “want to commit suicide,” “think about the future,” “lack of everything”
Murray et al., 2006	“worries,” “losing sleep,” “thinking too much,” “losing appetite,” “withdrawal,” “loneliness,” “suffering inside,” “feeling hopeless,” “wanting to commit suicide”

Table 3

Most frequently mentioned features across populations (N = 170) and in non-Western populations.

All populations (N = 178)		All non-Western populations (N = 115)	
Features	Frequency (%)	Features	Frequency (%)
Depressed mood/sadness^a	117 (68.8)	Depressed mood/sadness	78 (67.8)
Fatigue/loss of energy	100 (58.8)	Problems with sleep	73 (63.4)
Problems with sleep	99 (58.2)	Fatigue/loss of energy	73 (63.4)
Social isolation/loneliness	92 (54.1)	Social isolation/loneliness	60 (52.2)
Appetite/weight problems	75 (44.1)	Appetite/weight problems	58 (50.4)
Crying a lot	77 (43.5)	Crying a lot	53 (46.0)
Suicidal thoughts	69 (40.6)	Headaches	52 (45.2)
Loss of interest	68 (40.6)	Suicidal thoughts	50 (43.5)
Anger	64 (37.6)	General aches and pains	48 (41.7)
Headaches	60 (35.3)	Thinking too much	45 (39.1)
General aches and pains	58 (34.1)	Loss of interest	44 (38.2)
Worry	50 (29.4)	Issues with the heart ^b	42 (36.5)
Thinking too much	50 (29.4)	Anger	41 (35.6)
Issues with the heart ^b	50 (29.4)	Worry	38 (33.0)
Irritability & Worthlessness/guilt	49 (28.8)	Worthlessness/guilt	37 (32.2)

^a Bold indicates a diagnostic criterion for Major Depression in DSM-5.

^b Issues with the heart included features related to palpitations, feeling weakness in the heart, feeling as though heart was heavy, and heart pain.

Table 4

Top 5 most frequent features by region.^a

	Western non-Indigenous	Latin America	Middle East	East Asia	South Asia	Southeast Asia	Sub-Saharan Africa
1	Depressed mood ^b	Fatigue	Depressed mood	Fatigue	Sleep	Issues with heart ^{**}	Depressed mood
2	Social isolation/loneliness	Depressed Mood	Social isolation/loneliness	Worthlessness/guilt	Fatigue	Depressed mood	Sleep
3	Fatigue	Crying	Fatigue	Sleep	Depressed mood/sadness	Sleep	Weight/appetite
4	Sleep	Loss of interest	Irritability	Loss of interest	Weight/Appetite	Social isolation/loneliness	Fatigue
5	Loss of interest	Social isolation/loneliness	Sleep/General pain	Weight/Appetite & suicide	Issues with heart ^{**} /Headaches ^{**}	Thinking too much ^{**}	Social isolation/loneliness

^{*} $p < 0.05$.

^{**} $p < 0.005$ indicating statistically significant difference in frequency by region after corrected for multiple comparisons.

^a Northern America/Europe/Australia Indigenous populations not included in table because of the small number of studies reported from these regions.

^b Bold indicates a diagnostic criterion for Major Depression in DSM-5.

Table 5

Top 10 most frequently mentioned features among studies of single-gender populations.

Male only (<i>n</i> = 11)		Female only (<i>n</i> = 65)	
Features	Frequency (%)	Features	Frequency (%)
Depressed mood/sadness^a	8 (72.7)	Depressed mood/sadness	42 (64.6)
Angry	6 (54.5)	Fatigue	39 (60.0)
Social isolation/loneliness	6 (54.5)	Sleep	37 (56.9)
Weight/appetite	6 (54.5)	Social isolation/loneliness	33 (50.8)
Suicidal thoughts	6 (54.5)	Weight/appetite	25 (38.5)
Sleep	5 (45.5)	General pain and crying	23 (35.4)
Loss of interest	5 (45.5)	Crying	23 (35.4)
Irritability	5 (45.5)	Suicidal thoughts	23 (35.4)
Fatigue	5 (45.5)	Headaches	22 (33.8)
Crying, staying in bed, and thinking too much ^b	4 (36.4)	Heart issues and Irritability^c	19 (29.2)

* $p < 0.05$.

** $p < 0.005$ indicating statistically significant difference in frequency by gender or contextual variable after corrected for multiple comparisons.

^a Bold indicates included in DSM-5 diagnostic criteria for Major Depression.

^b All three features arose in 4 study populations.

^c Both features arose in 19 study populations.

Table 6

Top 10 most frequently mentioned features in the peri- and post-natal and trauma contexts.

Peri- and post-natal (<i>n</i> = 27)		Trauma (<i>n</i> = 25)	
Features	Frequency (%)	Features	Frequency (%)
Social isolation/loneliness	19 (70.4) *	Problems with sleep *	17 (68.0)
Depressed mood/sadness^a	18 (66.6)	Social isolation/loneliness *	16 (64.0)
Fatigue/lack of energy	15 (55.5)	Depressed mood/sadness	15 (60.0)
Problems with sleep	13 (48.2)	Appetite/weight *	15 (60.0)
Appetite/weight	11 (40.7)	Crying	14 (56.0)
Irritability	11 (40.7)	Hopelessness **	13 (52.0)
Crying	11 (40.7)	Suicidal thoughts	13 (52.0)
Anger	9 (33.3)	Fatigue/lack of energy	12 (48.0)
Suicidal thoughts	9 (33.3)	Function	11 (44.0)
Anxiety	9 (33.3)	Anger	11 (44.0)

* $p < 0.05$.

** $p < 0.005$ indicating statistically significant difference in frequency by gender or contextual variable after corrected for multiple comparisons.

^a Bold indicates included in DSM-5 diagnostic criteria for Major Depression.