BMC
Psychiatry

RESEARCH ARTICLE

Open Access

Comparing the accuracy of brief versus long depression screening instruments which have been validated in low and middle income countries: a systematic review

Dickens Akena[1,2*], John Joska[1], Ekwaro A Obuku[3], Taryn Amos[1], Seggane Musisi[2] and Dan J Stein[1]

Abstract

Background: Given the high prevalence of depression in primary health care (PHC), the use of screening instruments has been recommended. Both brief and long depression screening instruments have been validated in low and middle income countries (LMIC), including within HIV care settings. However, it remains unknown whether the brief instruments validated in LMIC are as accurate as the long ones.

Methods: We conducted a search of PUBMED, the COCHRANE library, AIDSLINE, and PSYCH-Info from their inception up to July 2011, for studies that validated depression screening instruments in LMIC. Data were extracted into tables and analyzed using RevMan 5.0 and STATA 11.2 for the presence of heterogeneity.

Results: Nineteen studies met our inclusion criteria. The reported prevalence of depression in LMIC ranged from 11.1 to 53%. The area under curve (AUC) scores of the validated instruments ranged from 0.69-0.99. Brief as well as long screening instruments showed acceptable accuracy (AUC≥0.7). Five of the 19 instruments were validated within HIV settings. There was statistically significant heterogeneity between the studies, and hence a meta-analysis could not be conducted to completion. Heterogeneity chi-squared = 189.23 (d.f. = 18) p<.001.

Conclusion: Brief depression screening instruments in both general and HIV-PHC are as accurate as the long ones. Brief scales may have an edge over the longer instruments since they can be administered in a much shorter time. However, because the ultra brief scales do not include the whole spectrum of depression symptoms including suicide, their use should be followed by a detailed diagnostic interview.

Background

Depression is a prevalent and disabling condition in both high and low income countries [1-3]. According to the World Health Organization, depression is the 4th most disabling medical disorder, and is predicted to be the 2nd most disabling medical condition by 2020 [1,4]. The 12-month prevalence of depression has been reported as 4.1%, with a lifetime prevalence of 6.7% [5].

Treatment guidelines developed in high income countries (HIC) recommend routine screening for depression in primary health care (PHC) as an initial step in holistic patient care [6-8]. A number of brief (≤12 items) instruments including the patient health questionnaire (PHQ-9) [9,10] and the Kessler-10 (K-10) [11] have been validated in low and middle income countries (LMIC). Similarly, longer (≥15 items) instruments including the centre for epidemiological studies-depression (CES-D) [12] have also been validated in LMIC.

The bulk of research summarizing findings about the accuracy of validated depression screening instruments has come from HIC, providing conflicting data [13-15]. For example, one review found marginal differences between brief and ultra-brief scales [14], while a meta-analysis by Mitchell et al. (2007) reported that brief and ultra-brief scales were equally accurate [15].

Generalizing findings from studies conducted in HIC to LMIC may be inappropriate due to a number of

* Correspondence: akenadickens@yahoo.co.uk
[1]Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa
[2]Department of Psychiatry, Makerere University, Kampala, Uganda
Full list of author information is available at the end of the article

differences. Low literacy rates, cultural diversity and high patient numbers are some factors that are unique to LMIC [3,16,17]. Such differences as low literacy rates may influence the accuracy of depression screening instruments, making the generalization of findings from HIC to LMIC the more difficult.

Depression is a major health problem across LMIC; however, a number of countries in sub-Saharan Africa are equally plagued with a high burden of HIV/AIDS. Indeed close to two thirds of all persons living with HIV/AIDS (PLWHA), reside in sub-Saharan Africa [18]. Research has also shown that up to 30% of PLWHA may develop depressive disorder during the course of their illness [19,20].

The screening of depression among PLWHA is important for a number of reasons; the presence of symptom overlap between the two disorders being one of them. For example, suicide, fatigue, sadness and insomnia are symptoms reported by both PLWHA and those with depression. The existence of symptom overlaps call for screening PLWHA who present at PHC for depression. Indeed a number of researchers have recommended the routine screening of depression in PLWHA [21-24]. However, literature about the validity of screening instruments in the setting of HIV/AIDS remains scanty [25].

The aim of our systematic review was to examine the accuracy of depression screening instruments which have been validated in LMIC, comparing brief and long scales. We also compared the accuracy of instruments validated in general and HIV-PHC settings.

These findings could guide clinicians about which scales to adapt for routine use in busy PHC settings within LMIC.

Methods

A literature search was conducted using the following approach:

We searched the PUBMED, COCHRANE library, AIDSLINE, and PSYCH-Info databases for studies published in English from inception up to July 2011. In our search, we used the following key words: sensitivity/specificity, validation, depression/depressive disorders, and screening instruments/tools/scales. These key words were combined with LMIC, HIV/AIDS, Africa, Asia, Eastern Europe, and South America. We then searched reference lists from retrieved articles for suitable papers and consulted two sets of authors [26,27] for more clarity regarding data in their papers.

Study selection

Studies were included if they had the following outcomes of interest:

1. A depression screening instrument followed by a formal diagnostic instrument or an interview was administered to all screened patients i.e. both screen positive and negatives.
   The diagnosis of a depressive disorder(major/minor/dysthymia) was based on the ICD-10 [28], DSM-IV [29], or an instrument frequently used as a gold standard. Instruments routinely used to screen for depression including the [30,31] were not considered gold standard, even though a number of studies had used them [25,32].
2. Studies were conducted in non-mental health facilities
3. Studies reported the sensitivity, specificity, the AUC and predictive values of the screening instrument in comparison to the diagnostic standard.
4. Studies were conducted in LMIC as defined by the world bank [33].

Data analysis

Data from included studies was extracted by one author (DA) into tables constructed in MS Excel, and later transferred to RevMan version 5.1.2 [34]. We used RevMan to construct a diagnostic 2x2 table by calculating the true positive, false positive, false negative and true negative figures from the sensitivity/specificity and prevalence values provided in all the included studies. The figures from the 2x2 tables generated using RevMan were then fitted in STATA version 11.2 [35] to assess for heterogeneity using random effects analysis model. Assessing for heterogeneity guided us, as to whether it was possible to pool, analyze, and report the findings as a meta-analysis. We used meta-analytic commands in STATA for the analysis.

Study quality assessment and inclusion

Data was independently abstracted by three authors (DA, EO and TA). DA read all the abstracts, 1151 studies were excluded based on abstracts alone. Full articles for 65 articles were identified for further scrutiny. Of the 65 articles identified for further scrutiny, 14 studies in which 19 instruments were validated with 3759 participants met our criteria. See Figure 1.

Study inclusion and exclusion was independently done by DA, EO and TA, in the event of ambiguity, DJS was the arbitrator. We used RevMan to assess study quality. The parameters assessed included blinding of reference information from screening results, screening of patients from highly selected populations, and selection of who gets the gold standard from among a screened population. Study quality was rated as fair, acceptable and good quality. All included studies were then scrutinized independently by JJ.

Figure 1 Study selection process for the systematic review.
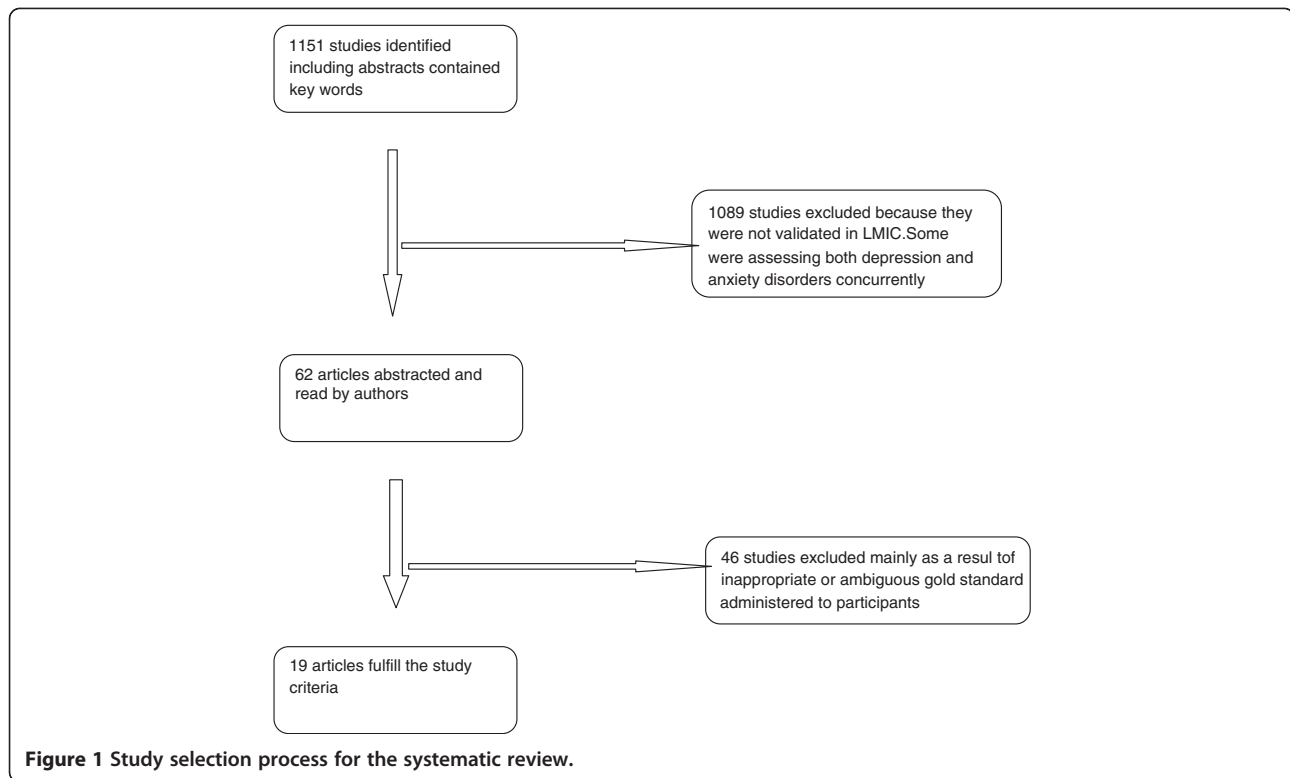
Results

Of the 19 included studies, 10 fulfilled all the reporting criteria by RevMan [30] and were considered of good quality [26,36-42].One study was considered fair in quality due to the lack of blinding and referral of only screen positives for the diagnosis from a highly selected population [11]. The rest of the studies (n=8) were considered acceptable. The studies with acceptable quality had limited information about blinding, some lacked clarity about the time interval between administration of the screening instrument and gold standard [27,43-47].

General description of studies

Eleven studies were conducted in Africa [11,26,27,38,40-43,47], five of which were in HIV settings [26,27,38,41,43]. Two studies were conducted in South America [36,37] and six in Asia [39,44-46] The most frequently used diagnostic instrument was the mini international neuropsychiatric instrument (MINI) [48]. Table 1 below shows the general characteristics of the studies. The sample sizes of included studies ranged from 61 to 649. The prevalence of depression varied widely across populations ranging from 11.1 to 53.5% (see Table 2 below). There were also wide variations within continents, and also according to the different instruments used. All validated instruments were able to adequately identify depression, with AUC ranging from 0.69-0.99. Table 2 above shows the variables that were used to assess for heterogeneity.

a) The BDI-SF, 1instrument

Leticia et al. (2005) [36] validated the BDI-SF validated among 155 patients admitted to general medical wards in Brazil. The gold standard was based on the ICD-10 [28].

b) K-6, 1 instrument.

Tesfaye et al. (2009) validated the K-6 in 100 post natal women attending a general PHC clinic in Ethiopia. A psychiatric interview based on the DSM-IV [29] was used as the gold standard.

c) K-10, 4 instruments

The K-10 was validated at four PHC sites, one of which was an HIV PHC site. Fernandes et al. (2011) [45] validated the K-10 among 194 pregnant mothers at a rural prenatal clinic in India. Meanwhile Spies et al. (2009) [27] validated the K-10 in 429 HIV-infected adults in an HIV care centre in South Africa using the MINI as the gold standard. Baggaley et al. (2007) [11] validated a translated version of the K-10 in Burkina Faso among 61 women. A detailed diagnostic interview by a psychiatrist within 3 days of administering the K10 was the gold standard. Tesfaye et al. (2009) validated the K-10 in 100 post natal women attending a general PHC clinic in Ethiopia. A psychiatric interview based on the DSM-IV [29] was used as the gold standard.

Table 1 General description of the studies included in the systematic review

Instrument	Author	Gold standard	Participant characteristics	Country of study
BDI-SF [a]	Leaticia et al. (2005)	ICD-10	Male/Female, in medical wards	Brazil
EPDS [b]	Chibanda et al. (2010)	MINI	Female, postnatal HIV-PHC	Zimbabwe
EPDS [b]	Lau et al. (2010)	SCID	Female, postnatal, general-PHC	China
EPDS [b]	Fernandes et al. (2011)	MINI	Female, antenatal, general-PHC	India
EPDS [a]	Figeuira et al. (2009)	MINI	Female, postnatal, general-PHC	Brazil
EPDS [b]	Tesfaye et al. (2009)	DSM-IV	Female,postnatal, general-PHC	Ethiopia
CESD [a]	Chisanga et al. (2011)	MINI	Male/Female HIV-PHC	Zambia
CESD [a]	Myer et al. (2008)	MINI	Male/Female HIV-PHC	South Africa
K-6 [a]	Tesfaye et al. (2009)	DSM-IV	Female,postnatal, general-PHC	Ethiopia
K-10 [c]	Baggeley et al. (2007)	Psychiatrist	Female, post natal, general-PHC	Burkina
K-10 [b]	Fernandes et al. (2011)	MINI	Female, antenatal, general-PHC	India
K-10 [b]	Spies et al. (2009)	MINI	Male/Female HIV-PHC	South Africa
K-10 [a]	Tesfaye et al. (2009)	DSM-IV	Female,postnatal, general-PHC	Ethiopia
PHQ-9 [a]	Lotrakul et al. (2008)	SCID	Male/Female Family practice clinic.	Thailand
SRQ-20 [a]	Stewart et al. (2009)	MINI	Female. Postnatal, General-PHC.	Malawi
VAS [b]	Puertas et al. (2004)	CIS-R	Male/Female General-PHC	India
GHQ-10 [b]	Puertas et al. (2004)	CIS-R	Male/Female General-PHC	India
SWB-4 [b]	Muwhezi et al. (2007)	MINI	Male/Female General-PHC	Uganda
HSCL-25 [a]	Kaaya et al. (2002)	SCID	Female, antenatal, HIV-PHC	Tanzania

*a denotes study of good quality, *b acceptable quality, and *c denotes study of fair quality.

d) PHQ-9, 1 instrument

The English language version of PHQ-9 was translated into Thai by Lotraku et al. (2008) [39], then back translated and adapted for use in Thailand. The PHQ-9 was then validated among 280 participants in a general PHC setting in Thailand.

e) EPDS, 5 instruments.

The EPDS was the most validated instrument in both pre and postnatal women. However, it should be noted that women accessing antenatal and postnatal care predominantly seek help for pregnancy related complaints, and may differ from persons attending general PHC. Despite such differences in the reason for seeking help at PHC, studies report a 10-20% prevalence of depression in postnatal women [49-51]. This high prevalence calls for the need to screen for depression in this population. We also report about these studies because such findings could be of interest to persons involved in women's mental health research.

Fernandes et al. (2011) [45] validated the EPDS among 194 women in their third trimester of pregnancy at a rural prenatal clinic in Karnataka India. The gold standard against which the EPDS was validated was the ICD-10. In mainland China, Lau et al. (2010) [44] validated the Chinese version of the EPDS in 342 postnatal women, using the Structured Clinical Interview for DSM-III-R (SCID) [52] as gold standard.

In Zimbabwe, Africa, Chibanda et al. (2010) [43] validated the Shona version of EPDS scale among 210 postpartum HIV-infected and uninfected women attending two primary care clinics in peri-urban Harare, Zimbabwe. In Brazil, Figeuira et al. (2009) [37] validated the EPDS in a sub-sample of 245 mothers; the MINI was used as the gold standard.

Tesfaye et al. (2009) validated the EPDS in 100 post natal women attending a general PHC clinic in Ethiopia. A psychiatric interview based on the DSM-IV [29] was used as the gold standard.

f) Other brief (3) instruments

Puertas et al. (2004) [46] validated a visual analogue scale (VAS) and the GHQ-10 among 450 participants in India using the revised Clinical Interview Schedule (CIS-R) [53] as a gold standard. The CIS-R is based on the ICD-10 [28].

In Uganda, Muwhezi et al. (2007) [47] assessed the validity of a 4- item subjective well-being subscale (SWB) in detecting a major depressive illness. A total of 199 consecutive patients were enrolled at a PHC facility in Uganda, interviewed using the SWB and the MINI [48] as a gold standard.

Longer scales

a) CES-D, 2 instruments

In Zambia, Africa, Chisanga et al. (2011) [38] conducted a cross-sectional study in 16 primary level

Table 2 Parameters used to asses for heterogeneity of included studies

Instrument	Author	No. of subjects	Prevalence	Sensitivity	Specificity	AUC
BDI-SF	Leaticia et al. (2005)	155	20	100	83.1	0.98
EPDS	Chibanda et al. (2010)	210	30.4	88	87	0.82
EPDS	Lau et al. (2010)	342	22.2	81.2	80.7	0.89
EPDS	Fernandes et al. (2011)	194	14.4	100	84.9	0.95
EPDS	Figeuira et al. (2009)	245	26.9	86.4	91.1	0.94
EPDS	Tesfaye et al. (2009)	100	11.0	78.9	75.3	0.85
CESD	Chisanga et al. (2011)	659	13.1	73	76	0.78
CESD	Myer et al. (2008)	465	13.3	79	61	0.75
K-6	Tesfaye et al. (2009)	100	11.0	82.4	82.7	0.86
K-10	Baggeley et al. (2007)	61	43.3	74	76	0.77
K-10	Fernandes et al. (2011)	194	14.4	100	81.3	0.95
K-10	Spies et al. (2009)	425	53.3	67	77	0.77
K-10	Tesfaye et al. (2009)	100	11.2	84.2	77.8	0.87
PHQ-9	Lotrakul et al. (2008)	280	6.78	84	77	0.89
SRQ-20	Stewart et al. (2009)	114	30.5	59.2	85.4	0.85
VAS	Puertas et al. (2004)	450	48.5	75.5	63.3	0.69
GHQ-10	Puertas et al. (2004)	450	48.5	93.6	81.1	0.87
SWB-4	Muwhezi et al. (2007)	199	37.3	75.7	86.3	0.87
HSCL-25	Kaaya et al. (2002)	100	11.1	89	80	0.86

care clinics and validated the CES-D in PLWHA who had tuberculosis and were starting ART. Chisanga validated the CES-D against the MINI [48] as gold standard.

Myer et al. (2008) [26] validated the CES-D among 465 participants individuals had enrolled into HIV care in South Africa. He used the MINI as gold standard.

b) SRQ-20, 1 instrument

In Malawi, Stewart et al. (2009) [40] validated the Chichewa version of the Self Reporting Questionnaire (SRQ) was validated among 114 subjects at a PHC site. This instrument went through a process of forward and back translation.

c) Other long instruments

Kaaya et al. (2002) [41] validated the Hopkins Symptom Checklist-25 (HSCL-25) among 99 women who were pregnant and HIV positive in Tanzania. The gold standard was the SCID [52].

Analysis for the presence of heterogeneity between studies

We used the 'meta' commands of STATA to generate the forest plots and assess for heterogeneity. The test for heterogeneity using a random effects analysis model yielded a statistically significant result. Heterogeneity chi-squared = 189.23, p = 0.000 on 18 degrees of freedom.

Statistically significant heterogeneity meant we could not continue with the meta-analysis and report the results as pooled estimates.

Discussion

We present the first systematic review comparing the accuracies of brief and long depression screening instruments which have been validated in LMIC settings. In this review, we found evidence to show that within LMIC, a number of depressed patients are identified using screening instruments at PHC settings. The prevalence figures reported in the included studies also vary widely across PHC settings within LMIC.

We found statistically significant heterogeneity between studies and could not conduct a meta-analysis to the end. The heterogeneity across studies could be the result of methodological differences in validation of instruments. For example, we found that a single instrument could be validated using different reference standards, producing different cut off scores and AUC scores. The CESD and EPDS were such examples in our review [26,38,43,45]. In addition, these studies were conducted across continents and settings with different cultures, languages and resources.

Both brief and longer scales showed moderate to high accuracy, with AUC ranging from 0.69-0.99. Our review found evidence to show that brief scales including the PHQ-9, BDI-SF, K-6, K-10, EPDS, and GHQ-12 were as accurate as the longer ones like the CES-D, HSCL, and BDI. These findings are in agreement with previous reviews which assessed the accuracy of depression screening instruments in HIC [6,14]. For example, a review of instruments validated in the Spanish language reported overall sensitivity and specificity in the range of

70-90% [13]. Studies with AUC's values of 0.50 to 0.70 are generally considered of low accuracy, 0.70 to 0.90 as having moderate accuracy, and those with AUC ≥ 0.90 as highly accurate [54,55]. Of the instruments studied, the EPDS shows acceptable accuracy in detecting depression among pre and post-natal women, which was in agreement with a previous systematic review [50]. Among HIV clinic populations, the HSCL-25 [41] showed the highest sensitivity at 89%.

No single instrument was superior to another in our review, perhaps due the relatively small number of studies with any particular instrument. Previous reviews that have assessed diagnostic accuracy of depression instruments were equally unable to recommend a single instrument for use in PHC [15,50].

Limitations

A number of limitations should be acknowledged. For example, we did not include studies that were not published in English. That said, our literature review did not return any studies in other languages that appeared to meet our inclusion criteria. While some studies published in non-indexed journals may have escaped notice, there has been an increase in indexed journals in LMIC in recent years, and most studies of quality should therefore have been captured.

Secondly, we didn't include in our review instruments which had been used to screen for the whole range of psychiatric morbidity, limiting our scope to those that had been validated for depression only. The inclusion of such scales which had screened for both depression and anxiety disorders could have been more informative; however, such criteria could have turned up numerous studies which may have been difficult to synthesize. Much as the K-10, GHQ and SRQ-20 instruments asses for common mental disorders including anxiety, depression and psychological distress, we only included them if they had been used to screen for depression.

Conclusion

Brief instruments are as accurate as the longer ones in detecting depression in both general and HIV-PHC settings. The brief nature of a screening instrument (BDI-SF, PHQ-10, and K-10) gives it the edge over longer scales like the CES-D due the short duration in which it can be administered. However, the fact that ultra-brief scales such the K-6 and BDI-SF don't encompass a whole range of depressive symptoms including suicide, the use of such scales needs to be followed up with detailed psychiatric diagnostic interviews. The K-6 was shown to be as accurate as the K-10 in the study by Tesfaye et al. (2009).

Other scales such as the EPDS may be the instrument of choice in particular populations (e.g. postnatal mothers).

Author details
[1]Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa. [2]Department of Psychiatry, Makerere University, Kampala, Uganda. [3]Joint Clinical Research Centre, Kampala, Uganda.

References
1. Murray CJL, Lopez AD: **Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study.** *Lancet* 1997, **349**:1436. 42 1997.
2. Martin P, Vikram P, Shekhar S, Mario M, Joanna M, Phillips MR, Atif R: **Global Mental Health 1 No health without mental health.** *Lancet* 2007, **370**:859. 77 2007.
3. Patel V: **Mental health in low- and middle-income countries.** *Br Med Bull* 2007, **81**(82):81–96.
4. World, Health, Organization: *Mental Health Report, managment of depression.* Geneva: World Health Organization; 2011. www.hoint/mental_health/management/depression/definition/en/.
5. Waraich P, Goldner EM, Somers JM, Hsu L: **Prevalence and incidence studies of mood disorders: a systematic review of the literature".** *Can J Psychiatry* 2004, **49**(2):124–138.
6. Gilbodyl S, Sheldon T, House A: **Screening and case-finding instruments for depression:a meta-analysis.** *Canadian Medical Journal* 2008, **178**(8):997–1003.
7. United States Preventive Services Task Force (USPSTF): **Screening for Depression,Recommendations and Rationale.** *Ann Intern Med* 1996, **136**(10):760–764.
8. National Institute of Clinical Excellence (NICE): *Management of Depression in Primary and Secondary Care.* London: NICE; 2004.
9. Adewuya AO, Ola BA, Afolabi OO: **Validity of the patient health questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students.** *J Affect Disord* 2006, **96**:89–93.
10. Shulin C, Helen C, Baihua X, Yan M, Tao J, Manhua W, Yeates C: **Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care.** *Int J Geriatr Psychiatry* 2010, **25**:1127–1133.
11. Baggaley RF, Ganaba R, Filippi V, Kere M, Marshall T, Sombie I, Storeng KT, Patel V: **Short communication: Detecting depression after pregnancy the validity of the K10 and K6 in Burkina Faso.** *Trop Med Int Health* 2007, **12**(10):1225–1229.
12. Sheung-Tak C, Chan ACM: **The Center for Epidemiologic Studies Depression Scale in older Chinese: thresholds for long and short forms.** *Int J Geriatr Psychiatry* 2005, **20**:465–470.
13. Reuland DS, Cherrington A, Watkins GS, Bradford DW, Blanco RA, Gaynes BN: **Diagnostic Accuracy of Spanish Language Depression-Screening Instruments.** *Ann Fam Med* 2009, **7**:455–462.
14. Henkel V, Mergl R, Kohnen R: **Use of brief depression screening tools in primary care: consideration of heterogeneity in performance in different patient groups.** *Gen Hosp Psychiatry* 2004, **26**:190–198.
15. Mitchell AJ, Coyne JC: **Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies.** *Br J Gen Pract* 2007, **57**(535):144–151.
16. Patel V, Simon G, Chowdhary N, Kaaya S, Araya R: **Packages of Care for Depression in Low- and Middle-Income Countries.** *PLoS Med* 2009, **6**(10):e1000159.

17. Vikram P, Weiss HA, Neerja C, Smita N, Sulochana P, Sudipto C, De Silva MJ, Bhargav B, Ricardo A, Michael K, et al: **Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): a cluster randomised controlled trial.** *Lancet* 2010, **376**:2086–2095.

18. World Health Organization: **HIV/AIDS and mental health.** In *EB124/6*. Edited by Executive Board 124th Session Provisional agenda item 4.3. Geneva: World Health Organization; 2008.

19. Nakimuli-Mpungu E, Bass JK, Alexandre P, Mills EJ, Musisi S, Ram M, Katabira E, Nachega JB: **Depression, Alcohol Use and Adherence to Antiretroviral Therapy in Sub-Saharan Africa: A Systematic Review.** *AIDS Behav* 2011, **15**:376–388.

20. Ciesla JA, Roberts JE: **Meta-Analysis of the Relationship Between HIV Infection and Risk for Depressive Disorders.** *Am J Psychiatry* 2001, **158**:725–730.

21. Nakimuli-Mpungu E, Mutamba B, Othengo M, Musisi S: **Psychological distress and adherence to highly active anti-retroviral therapy (HAART) in Uganda: A pilot study.** *Afr Health Sci* 2009, **9**(S2):2–7.

22. Cook JA, Dennis G, Jane B, Cohen MH, Gurtman AC, Richardson JL, Wilson TE, Young MA, Hessol NA: **Depressive Symptoms and AIDS-Related Mortality Among a Multisite Cohort of HIV-Positive Women.** *Am J Public Health* 2004, **94**:1133–1140.

23. Noeline N, Skolasky RL, Seggane M, Peter A, Kevin R, Allan R, Elly K, Clifford DB, Ned S: **Depression symptoms and cognitive function among individuals with advanced HIV infection initiating HAART in Uganda.** *BMC Psychiatry* 2010, **10**:44.

24. Olley BO, Seedat S, Nel DG, Stein DJ: **Predictors of Major Depression in Recently Diagnosed Patients with HIV/AIDS in South Africa.** *AIDS Patient Care STDS* 2004, **18**:481–487.

25. Monahan PO, Enbal S, Michael R, Kurt K, Willis Owino O'o, Otieno O, Violet Naanyu Y, Claris O: **Validity/Reliability of PHQ-9 and PHQ-2 Depression Scales Among Adults Living with HIV/AIDS in Western Kenya.** *J Gen Intern Med* 2008, **24**(2):189–197.

26. Landon M, Joalida S, Le Liezel R, Siraaj P, Stein DJ, Soraya S: **Common Mental Disorders among HIV-Infected Individuals in South Africa: Prevalence, Predictors, and Validation of Brief Psychiatric Rating Scales.** *AIDS Patient Care STDS* 2008, **22**(2):147–158. doi:10.1089/apc.2007.0102.

27. Gordon S, Adul K, Kidd M, Smit J, Landon M, Dan S, Soraya S: **Validity of the K-10 in detecting DSM-IV-defined depression and anxiety disorders among HIV-infected individuals.** *AIDS Care* 2009, **21**:1163–1168.

28. World Health Organisation: *ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization; 1992.

29. American Psychiatric Association: *Diagnostic and statistical manual of mental disorders (4th ed., text rev.)*. 2000.

30. Radloff LS: **Center for Epidemiologic Studies Depression Scale (CESD).** In *National Institute of Mental Health, Centre for Epidemiologic Studies*. USA: West Publishing Company; 1977.

31. Spitzer R, Kroenke K, Williams J: **Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study.** *J Am Med Assoc* 1999, **282**:1737–1744.

32. Emilio O, Jed B, Danuta W: **The Response Inventory for Stressful Life Events (RISLE) I. refinement of the 100-item Version.** *Afr Health Sci* 2005, **5**(2):137–144.

33. World Bank: *Classification of countries according to income levels, webpage accessed on the 10th July 2011*. Washington DC: World Bank; 2011. www.data.worldbank.org/about/country-classifications.

34. *The cochrane IMS*. Copenhagen: The Cochrane collaboration; 2008. www.ims.cochrane.org/revman/5.1.2 download.

35. STATA, StataCorp LP: *Statistics/Data Analysis, 4905 Lakeway Drive College Station, Texas 77845 USA 800-STATA-PC*. Texas: STATA; 2011. http://www.stata.com.

36. Furlanettoa LM, Mendlowicz MV, Buenob RJ: **The validity of the Beck Depression Inventory-Short Form as a screening and diagnostic instrument for moderate and severe depression in medical inpatients.** *J Affect Disord* 2005, **86**:87–91.

37. Patrícia F, Humberto C, Leandro M-D, Marco Aurélio R-S: **Edinburgh Postnatal Depression Scale for screening in the public health system.** *Rev Saude Publica* 2009, **43**(1):1–5.

38. Nathaniel C, Eugene K, Weiss HA, Vikram P, Helen A, Soraya S: **Validation of brief screening tools for depressive and alcohol use disorders among TB and HIV patients in primary care in Zambia.** *BMC Psychiatry* 2011, **11**:75.

39. Manote L, Sutida S, Ratana S: **Reliability and validity of the Thai version of the PHQ-9.** *BMC Psychiatry* 2008, **8**:46.

40. Stewart RC, Felix K, Eric U, Maclean V, James B, Margaret F, Barbara T, Atif R, Francis C: **Validation of a Chichewa version of the Self-Reporting Questionnaire (SRQ) as a brief screening measure for maternal depressive disorder in Malawi, Africa.** *J Affect Disord* 2009, **112**:126–134.

41. Kaaya SF, Fawzi MCS, Mbwambo K, Lee B, Msamanga GI, Fawzi W: **Validity of the Hopkins Symptom Checklist-25 amongst HIV-positive pregnant women in Tanzania.** *Acta Psychiatra Scandanivica* 2002, **106**:9–19.

42. Markos T, Charlotte H, Dawit W, Atalay A: **Detecting postnatal common mental disorders in Addis Ababa, Ethiopia:Validation of the Edinburgh Postnatal Depression Scale and Kessler Scales.** *J Affect Disord* 2009, **102**:8. doi:10.1016/j.jad.2009.1006.1020.

43. Chibanda D, Mangezi W, Tshimanga M, Woelk G, Rusakaniko P, Stranix-Chibanda L, Midzi S, Maldonado Y, Shetty AK: **Validation of the Edinburgh Postnatal Depression Scale among women in a high HIV prevalence area in urban Zimbabwe.** *Archives of Womens Mental Health* 2010, **13**:201–206.

44. Ying L, Yuqiong W, Lei Y, Kin Sun C, Xiujing G: **Validation of the Mainland Chinese version of the Edinburgh Postnatal Depression Scale in Chengdu mothers.** *Int J Nurs Stud* 2010, **47**:1139–1151.

45. Michelle Caroline F, Krishnamachari S, Stein AL, Gladys M, Sumithra RS, Ramchandani PG: **Assessing prenatal depression in the rural developing world: a comparison of two screening measures.** *Archives of Womens Mental Health* 2011, **14**:209–216.

46. Puertas G, Patel V, Marshall T: **Are visual measures of mood superior to questionnaire measures in non-Western settings?** *Social Psychiatry Epidemiology* 2004, **39**:662–666.

47. Wilson Winstons M, Hans A, Seggane M: **Detection of major depression in Ugandan primary health care settings using simple questions from a subjective well-being (SWB) subscale.** *Soc Psychiatry Psychiatr Epidemiol* 2007, **42**:61–69.

48. Sheehan DV, Lecrubier Y, Harnett-Sheehan K: **The Mini International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview.** *J Clin Psychiatry* 1998, **20**:22–33.

49. Breese MC, Sarah J: **Postpartum depression: an essential overview for the practitioner.** *South Med J* 2011, **104**(2):128–132.

50. Zubaran C, Schumacher M, Roxo MR, Foresti K: **Screening tools for postpartum depression: validity and cultural dimensions.** *Afr J Psychiatr* 2010, **13**:357–365.

51. Villegas L, McKay K, Dennis C-L, Ross LE: **Postpartum Depression Among Rural Women From Developed and Developing Countries: A Systematic Review.** *J Rural Health* 2011, **27**(2011):278–288.

52. Spitzer RL, Williams JBW, Miriam G, First MB: **The Structured Clinical Interview for DSM-III-R (SCID)I: History, Rationale, and Description.** *Arch Gen Psychiatry* 1992, **49**(8):624–629.

53. Glyn L, Pelosi AJ, Ricardo A, Graham D: **Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers.** *Psychol Med* 1992, **22**:465–486. doi:410.1017/S0033291700030415.

54. Swets JA: **Measuring the accuracy of diagnostic systems.** *Science* 1998, **240**(4857):1285–1293.

55. Fischer JE, Bachmann LM, Jaeschke R: **A readers' guide to the interpretation of diagnostic test properties.** *Clinical example of sepsis Intensive Care Medicine* 2003, **29**(7):1043–1051.