Validity/Reliability of PHQ-9 and PHQ-2 Depression Scales Among Adults Living with HIV/AIDS in Western Kenya

Patrick O. Monahan, PhD[1], Enbal Shacham, PhD[2], Michael Reece, PhD, MPH[3], Kurt Kroenke, MD[4], Willis Owino Ong'or, MMEd, MPH[6], Otieno Omollo, MBChB, MMed[6], Violet Naanyu Yebei, MA[5,6], and Claris Ojwang[7]

[1]Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA; [2]Department of Psychiatry, School of Medicine, Washington University, St. Louis, MO, USA; [3]Department of Applied Health Science, Indiana University, Bloomington, IN, USA; [4]Regenstrief Institute for Health Care and Department of Medicine, Indiana University, Indianapolis, IN, USA; [5]Department of Sociology, Indiana University, Bloomington, IN, USA; [6]School of Medicine, Moi University, Eldoret, Kenya; [7]AMPATH Support Network, Moi University, Eldoret, Kenya.

BACKGROUND: Depression greatly burdens sub-Saharan Africa, especially populations living with HIV/AIDS, for whom few validated depression scales exist. Patient Health Questionnaire-9 (PHQ-9), a brief dual-purpose instrument yielding DSM-IV diagnoses and severity, and PHQ-2, an ultra-brief screening tool, offer advantages in resource-constrained settings.

OBJECTIVE: To assess the validity/reliability of PHQ-9 and PHQ-2.

DESIGN: Observational, two occasions 7 days apart.

PARTICIPANTS: A total of 347 patients attending psychosocial support groups.

MEASUREMENTS: Demographics, PHQ-9, PHQ-2, general health perception rating and CD4 count.

RESULTS: Rates for PHQ-9 DSM-IV major depressive disorder (MDD), other depressive disorder (ODD) and any depressive disorder were 13%, 21% and 34%. Depression was associated with female gender, but not CD4. Construct validity was supported by: (1) a strong association between PHQ-9 and general health rating, (2) a single major factor with loadings exceeding 0.50, (3) item-total correlations exceeding 0.37 and (4) a pattern of item means similar to US validation studies. Four focus groups indicated culturally relevant content validity and minor modifications to the PHQ-9 instructions. Coefficient alpha was 0.78. Test-retest reliability was acceptable: (1) intra-class correlation 0.59 for PHQ-9 total score, (2) kappas 0.24, 0.25 and 0.38 for PHQ-9 MDD, ODD and any depressive disorder and (3) weighted kappa 0.53 for PHQ-9 depression severity categories. PHQ-2 $\geq$3 demonstrated high sensitivity (85%) and specificity (95%) for diagnosing any PHQ-9 depressive disorder (AUC, 0.97), and 91% and 77%, respectively, for diagnosing PHQ-9 MDD (AUC, 0.91). Psychometrics were also good within four gender/age (18–35, 36–61) subgroups.

CONCLUSIONS: PHQ-9 and PHQ-2 appear valid/reliable for assessing DSM-IV depressive disorders and depression severity among adults living with HIV/AIDS in western Kenya.

KEY WORDS: HIV/AIDS; Kenya; Africa; depression; PHQ-9.
J Gen Intern Med 24(2):189–97
DOI: 10.1007/s11606-008-0846-z

Received November 14, 2007
Revised June 30, 2008
Accepted October 8, 2008
Published online November 20, 2008

INTRODUCTION

More than 70% of all deaths attributable to HIV/AIDS are in sub-Saharan Africa, with approximately 6.1% of adults living with HIV in Kenya in 2005[1]. The social, economic and health impact of depression in sub-Saharan Africa is also great, where depression is associated with mortality[2–5], work disability[4–7], lower quality of life[5,8–12], risk of heart disease[13] and high-risk behaviors for contracting HIV infection[14]. With one exception[15], the sparse literature on depression among individuals living with HIV/AIDS in sub-Saharan Africa has shown elevated rates of depression relative to community samples[3,8,10,16–25], consistent with western countries[26–29]. Thus, there is an urgency to incorporating mental health into HIV/AIDS treatment programs in sub-Saharan Africa, including western Kenya[30–35].

Depression in sub-Saharan Africa presents in forms (culture-specific idioms, somatic, based on interpersonal relationships or spiritual in nature) that may obscure detection[36–38]. However, depression exists at possibly higher prevalence rates than in western countries[37] according to 16 studies[4,5,7,9–11,38–47], and an additional 13 studies cited in[37], of community and non-HIV-specific clinic populations, with generally higher rates for women than men[37,43]. Finally, depression is reasonably easy to elicit when sought and present[37].

Few studies have validated depression instruments for sub-Saharan cultures[10,11,46–52]. The Patient Health Questionnaire-9 (PHQ-9) is the only validated instrument[11,46] that focuses on the nine diagnostic criteria for DSM-IV depressive disorders and is very brief[53], an advantage in resource-constrained Kenyan clinics. The PHQ-9 can be self-administered[53,54] or interviewer-administered[11,55,56], and is well validated in the US as a dual-purpose instrument that yields both a measure of depression *severity*[57,58] and criteria-based *diagnoses* of DSM-IV depressive disorders[53,54]: *major depressive disorder*

(MDD), *other depressive disorder* (ODD) and *any depressive disorder* (i.e., MDD or ODD).

The PHQ-9 has been validated in western countries regarding construct validity[53,54,57,59–64], diagnostic accuracy [53,57,58,62,64–68], sensitivity to change[61,69,70], responsiveness to treatment[71–73], internal consistency[56,57,63,65,74,75], test-retest reliability[56,57,62,73] and realistic estimates of population base rates [74]. The PHQ-2 (i.e., the first two items of the PHQ-9) is an ultra-brief and accurate screening tool[58,76–78].

The PHQ-9 has been validated in chronically ill populations[58,62,63,79]. However, only one report studied individuals living with HIV (in the US); the PHQ-9 demonstrated better diagnostic accuracy than the provider report[27]. Only three published studies administered the PHQ-9 in sub-Saharan Africa, showing good psychometric properties in urban Kenyan low-literacy cancer patients[11], Nigerian university students[46] and educated Nigerian army personnel[80].

The present study addresses two questions in the context of adults living with HIV/AIDS in western Kenya:

1. What is the validity and reliability of the PHQ-9 for assessing DSM-IV depressive disorders and depression severity?
2. What are the operating characteristics of the PHQ-2 as a diagnostic screening tool for determining PHQ-9 DSM-IV diagnoses of MDD and any depressive disorder?

METHODS

Participant Recruitment

The protocol was approved by the Institutional Research and Ethics Committee at Moi University (Kenya) and the Committee on Protection of Human Subjects at Indiana University-Bloomington (US). Of those attending routine psychosocial support groups during the study week at Moi University Teaching and Referral Hospital in Eldoret, Kenya, 100% agreed to participate when invited by a research assistant[32]. These support groups are a component of a comprehensive HIV prevention and care program known as the Academic Model for the Prevention and Treatment of HIV/AIDS (AMPATH)[81–84].

Data Collection

An informed consent statement was read to individuals by research assistants in English (or Swahili if preferred) and signed by a "witness" (research assistant) to protect anonymity. Participants completed a self-administered paper-based baseline questionnaire in English (n=397). Research assistants were available to translate unfamiliar terms using an unpublished Swahili version our team developed through translation and back translation. Every other participant was invited, excluding focus group participants, to complete a retest assessment 1 week later (n=187). Analyses were based on 347 participants who had no missing baseline PHQ-9 items. Test-retest analyses included 145 participants who had no missing baseline or retest PHQ-9 items. Focus groups were conducted during the first week among baseline participants. The most recent CD4 count within the previous 6 months was abstracted from medical charts.

Measures

Depression was assessed with the PHQ-9 (Appendix)[53]. The PHQ-9 scoring algorithm for DSM-IV diagnoses of MDD and ODD is described elsewhere (p. 607)[57]. The PHQ-9 total score ranges from 0 to 27 with five severity categories: minimal (0–4), mild (5–9), moderate (10–14), moderately severe (15–19) and severe (20–27).

General health perception rating was administered[85]: "In general how would you rate your overall health right now?" Ratings were from excellent (1) to poor (5). Recommended scoring was used: 1=5, 2=4.36, 3=3.43, 4=1.99 and 5=1, rescaled 0–100; a higher score represents better perceived health[85].

Statistical Analysis

All analyses were performed with SPSS, except weighted kappa was computed with SAS. Significance level was 0.05.

Construct validity. A significant relationship between the PHQ-9 and general health rating[62,63] or functional impairment scales[53,54,57,60,64,76] supports construct validity, because depression is known to negatively impact perceived health[60]. Analysis of covariance (ANCOVA) was used to compare PHQ-9 diagnostic and severity categories (independent variable) on general health rating (dependent variable), while adjusting, consistent with previous studies[53,54,60], for gender, age and education (≥ Form 1–2 vs ≤ Standard 4–8). The false discovery rate method controlled the proportion (0.05) of falsely rejected null hypotheses for multiple post-hoc pair-wise comparisons of depression categories[86,87]. Results were descriptively compared to two large, demographically diverse, US validation samples[53,54].

Construct validity. *Factor analysis* was performed using principal component extraction, initial communalities of 1.0 and Varimax rotation. Number of factors was determined with the scree plot[88]. Item discrimination was assessed with corrected item-total correlations[89].

Content validity. Narrative data were collected from four single-gender[90] focus groups using cognitive interviewing[91]. Sample sizes were four and seven for female groups, and five and eight for male groups. There were two facilitators in each group (one American, one Kenyan). Participants could answer in Swahili when unsure of English terminology. Discussions examined meaning, comprehension and interpretation of PHQ-9, and lasted an average of 1.5 h.

Reliability. Internal consistency was assessed with coefficient alpha. Test-retest reliability was assessed for (1) PHQ-9 total score with an intraclass correlation coefficient (ICC) using agreement and consistency indices with occasions as

random[92], (2) PHQ-9 DSM-IV diagnoses with a kappa coefficient of chance-corrected agreement[93] and (3) PHQ-9 severity categories with weighted kappa[94] using Fleiss-Cohen quadratic weights[95].

Operating characteristics of the PHQ-2. The PHQ-2 operating characteristics were calculated for PHQ-9 MDD and separately for any PHQ-9 depressive disorder. Area under receiver-operator curve (AUC) and its nonparametric 95% CI were computed[96].

RESULTS

Participant Characteristics

The mean age was 36 years, 73% were female, and 38% were married (Table 1). Education levels completed were predominantly "Standard 4–8" (38%) and "Form 3–4" (34%), comparable to American elementary and high school, respectively. The majority of participants were unemployed but looking for work (69%). Median time living with HIV was 18 months; 44% reported being diagnosed with AIDS. Median CD4 count was 310 cells/μL (<50 cells/μL for 11 participants). Compared to males, females were significantly younger and less likely to be married, employed or diagnosed with AIDS, and had a lower CD4 count.

Prevalence of Depression Diagnoses and Severity

One-third of participants met PHQ-9 DSM-IV criteria for either MDD (13%) or ODD (21%). Fifty-three percent displayed mild or moderate depression severity (Table 1). Floor effects for the PHQ-9 total score were negligible (only 10.4% scored zero). There was a significant interaction between CD4 count and gender on the PHQ-9 total score (p=0.048). Females had significantly higher depression, but only among participants with a CD4 count <200 (Table 1, p=0.02). More importantly, after using ANCOVA to adjust for potentially confounding variables for which gender differed at least marginally (p<.10) in Table 1 (age; dichotomizations of marital status, education, employment and CD4 count; and AIDS diagnosis), the PHQ-9 score (dependent variable) was significantly (p=0.04) higher for females than males (adjusted means, 7.3 vs 5.6) in the entire sample regardless of the CD4 count, and the interaction was not significant.

Depression and HIV Progression

Regarding the interaction above, females with CD4 count <200 had a trend for a higher PHQ-9 total score than females with CD4 count ≥200 (p=0.14), and this trend was reversed for males (p=0.17). To investigate this counterintuitive result for males, a result consistent with our report on psychological distress[30], we accounted for confounding covariates by using the ANCOVA model in the previous paragraph after removing the AIDS diagnosis (because it was correlated with the CD4 count); the interaction was not significant, and the adjusted mean PHQ-9 score was 7.0 and 6.7 for a CD4 count <200 and ≥200, respectively (p=0.78), revealing very little relationship between the CD4 count and depression, but a trend in the anticipated direction (e.g.,[97]) for the entire sample regardless of gender.

Construct Validity: Relationship Between PHQ-9 and General Health Rating

PHQ-9 diagnostic categories differed significantly on general health rating (p=0.001). General health rating was best for no depressive disorder (adjusted mean, 63.3), lower for ODD (54.6) and worst for MDD (46.5). MDD (p<.0001) and ODD (p=0.029) were each significantly different from no depressive disorder. However, MDD did not differ significantly from ODD.

PHQ-9 severity categories also differed significantly according to general health rating (p<.0001). Minimal depression severity differed significantly from all other depression severity categories with respect to general health rating. Those with severe depressive symptoms had significantly lower general health perceptions than those with mild depressive symptoms. General health rating steadily decreased as depression severity increased, a relationship that was not markedly different from that observed in two US validation samples (Fig. 1). (For consistency, US samples excluded participants missing any PHQ-9 items.) Importantly, results did not vary significantly by age, gender or education (i.e., no significant interactions).

Construct Validity: Factor Analysis

The scree plot indicated one dominant dimension with a large decrease between first and second eigenvalues and small decreases thereafter (eigenvalues: 3.3, 1.0, 0.9, 0.8, 0.7, 0.7, 0.6, 0.5 and 0.5). Factor loadings ranged from 0.52 to 0.66 (i.e., above 0.40 cutoff)[98]. Factor loadings were strong (>0.50) for four subgroup analyses: females (Table 2, Part B), males (Table 2, Part C) and age groups dichotomized at median (age 18–35 and 36–61, not tabled). The percentage of total variance explained by the first factor was 37% (34%–40% for subgroups). Item-total correlations exceeded 0.37 (i.e., above the 0.30 cutoff)[98] in the total sample and four subgroups.

Low energy was the most highly endorsed item (mean, 1.10), as it was in US validation samples (Fig. 2). All items except sleep problems were more highly endorsed in the western Kenya sample compared to US samples. The pattern of means across items was not substantially different from US samples (Fig. 2).

Content Validity

Focus groups revealed that the PHQ-9 was generally well understood, acceptable and culturally relevant for all gender and age groups. Three findings emerged, applicable to all gender and age groups. First, interpreting opposite symptoms in Items 3, 5 and 8 was challenging for the majority of participants who repeatedly asked "Do you want both answers?" After discussion, participants understood the items referred to changes in behaviors in *either* direction. Second, participants understood the item scale and the 2-week recall period, but found it somewhat confusing to evaluate the item scale (e.g., "more than half the days") *in relation* to 2 weeks. Third, not surprisingly, participants reported interpreting and responding to PHQ-9 items mostly in the context of living with HIV. For example, Item 9 was perceived by both genders as the item referring most directly to living with HIV. Item 6 was perceived by males in the context of feeling guilty about getting infected and infecting his family with HIV.

Table 1. Participant Characteristics

	Total sample (n=347)		Females (n=251)		Males (n=93)		p*
	M (SD) Range or n (%)	nm†	M (SD) Range or n (%)	nm†	M (SD) Range or n (%)	nm†	
Sociodemographics							
Age	36.3 (7.9) 18–61	4	35.7 (8.2) 18–61	3	38.1 (6.7) 23–56	1	0.01
Number of children	3.1 (2.1) 0–11	2	3.1 (2.0) 0–9	1	3.2 (2.3) 0–11	1	0.78
Children live with you, yes	293 (90%)	23	217 (92%)	14	74 (88%)	9	0.38
Marital status‡		5		3		2	<.001
Married	131 (38%)		70 (28%)		59 (65%)		
Single	82 (24%)		71 (29%)		10 (11%)		
Widow/er	93 (27%)		74 (30%)		19 (21%)		
Divorced	36 (11%)		33 (13%)		3 (3%)		
Tribal affiliation§		6		4		2	0.17
Luyha	82 (24%)		65 (26%)		16 (18%)		
Kikuyu	80 (24%)		62 (25%)		18 (20%)		
Luo	62 (18%)		39 (16%)		22 (24%)		
Kalenjin	44 (13%)		30 (12%)		14 (15%)		
Nandi	21 (6%)		17 (7%)		4 (4%)		
Other	52 (15%)		34 (14%)		17 (19%)		
Religious affiliation		4		3		1	0.43
Protestant	203 (59%)		147 (59%)		53 (58%)		
Catholic	95 (28%)		66 (27%)		29 (32%)		
Muslim	12 (3%)		11 (4%)		1 (1%)		
Other	33 (10%)		24 (10%)		9 (10%)		
Education level		0		0		0	0.19
None	17 (5%)		15 (6%)		2 (2%)		
Standard 1–3	24 (7%)		21 (8%)		3 (3%)		
Standard 4–8	134 (38%)		98 (39%)		35 (38%)		
Form 1–2	48 (14%)		35 (14%)		13 (14%)		
Form 3–4	118 (34%)		77 (31%)		39 (42%)		
University	6 (2%)		5 (2%)		1 (1%)		
Employment status		29		20		9	0.03
Full time	37 (12%)		23 (10%)		14 (17%)		
Part time	28 (9%)		20 (9%)		7 (8%)		
Unemployed, looking	220 (69%)		163 (71%)		55 (65%)		
Unemployed, not looking	22 (7%)		20 (9%)		2 (2%)		
Other	11 (3%)		5 (2%)		6 (7%)		
Clinical							
Length time HIV, months‖	18 (27) 0–168	32	19 (25) 0–141	29	18 (29) 1–168	3	0.64
AIDS diagnosis	145 (44%)	17	96 (41%)	14	49 (54%)	3	0.02
Length time AIDS, months‖	12 (19) 0–180	203	15 (24) 0–180	157	12 (18) 1–90	43	0.10
CD4 count, cells/μL‖	310 (272) 2–1321	45	329 (262) 2–1321	33	222 (254) 7–1113	12	<.001
CD4 count <200 cells/μL	85 (28%)	45	47 (22%)	33	38 (47%)	12	<.001
Depression							
PHQ-9 total score	7.2 (5.5) 0–23	0	7.4 (5.4) 0–23	0	6.5 (5.6) 0–22	0	0.17
When CD4 count <200	6.7 (5.7) 0–21		8.1 (6.0) 0–21		5.1 (5.1) 0–18		0.02
When CD4 count ≥200	6.9 (5.3) 0–23		6.8 (5.1) 0–23		6.8 (5.7) 0–20		0.97
PHQ-9 DSM-IV diagnoses		0		0		0	0.93
Major depressive disorder	45 (13%)		32 (13%)		12 (13%)		
Other depressive disorder	73 (21%)		52 (21%)		21 (23%)		
PHQ-9 severity levels		0		0		0	0.56
Minimal (0–4)	124 (36%)		84 (34%)		39 (42%)		
Mild (5–9)	116 (33%)		90 (36%)		26 (28%)		
Moderate (10–14)	68 (20%)		48 (19%)		19 (20%)		
Moderately severe (15–19)	27 (8%)		21 (8%)		6 (7%)		
Severe (20–27)	12 (3%)		8 (3%)		3 (3%)		

*p value is from the comparison between males and females using the two-sided t test for age and number children; two-sided Wilcoxon rank sum test for length of time with HIV, length of time with AIDS and CD4 count; Pearson Chi-square test for categorical variables.

†nm = number of missing observations for the particular variable.

‡Married category included two persons not married but having a significant other/partner.

§Other category for tribal affiliation included (<3% from any single tribe): Baluyha, Busuku, Giriama, Kamba, Keiyo, Kipsigis, Kisii, Maasai, Meru, Muluya, Munandi, Munyakare, Sudanise, Teso, Tiriki, Tungen and Turkana.

‖ For positively skewed variables, the median (interquartile range) and range were reported instead of the mean (SD) and range. The p values for bivariate gender comparisons, in Table 1, of dichotomized versions of variables used in subsequent models were: married vs. not married, p<.001; ≥ Form 1–2 vs. ≤ Standard 4–8, p=.09; unemployed vs. employed and other, p=.04.

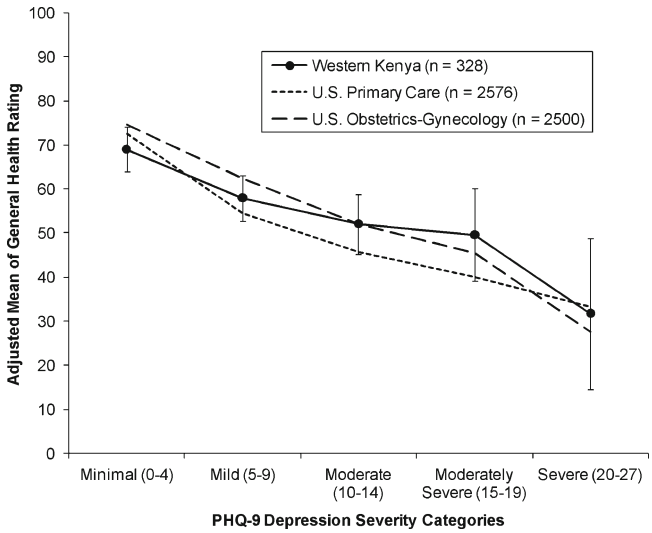

Figure 1. Relationship between PHQ-9 depression severity and general health rating. Vertical bars represent 95% confidence intervals of the adjusted means for the western Kenyan sample.

Reliability

Coefficient alpha for PHQ-9 total score was 0.78 at baseline (0.77, females; 0.81, males; 0.75, age 18–35; 0.81, age 36–61). One-week test-retest reliability of PHQ-9 total score was 0.59 (0.58, females; 0.60, males; 0.65, age 18–35; and 0.51, age 36–61) for both agreement and consistency ICC indices. The mean PHQ-9 total score decreased slightly, but not significantly over the two occasions (6.70 to 6.23, two-sided paired t test, p= 0.23). Kappa was 0.24 for MDD (p=0.003), 0.25 for ODD (p= 0.003) and 0.38 for any depressive disorder (p<.0001). Weighted kappa was 0.53 for PHQ-9 severity categories (p<.0001).

Operating Characteristics of the PHQ-2

The cutoff point of PHQ-2 ≥3 demonstrated high sensitivity (91.1%) and moderate specificity (76.8%) for PHQ-9 MDD (Table 3). PHQ-2 ≥4 showed high specificity (93.4%) but low sensitivity (57.8%). Those with MDD were 3.93 times more likely to have a PHQ-2 score ≥3 than those without MDD.

For any depressive disorder, PHQ-2 ≥3 was the best cutoff point for achieving high sensitivity (84.7%) and high specificity (95.2%); the positive predictive value was also high (90.1%) (Table 3). AUC was very high for MDD (0.91) and any depressive disorder (0.97) (Table 3).

DISCUSSION

Rates of depressive disorders were consistent with studies in sub-Saharan Africa. The PHQ-9 DSM-IV diagnostic and severity categories differed on general health rating, supporting construct validity. With a larger sample size, the difference between MDD and ODD may become significant, because their mean difference on general health rating was similar to the mean difference between ODD and no depressive disorder. The strong relationship between increasing PHQ-9 depression severity and worsening general health rating was consistent with two large US validation samples (see also [60]), supporting construct validity. Factor analysis revealed acceptably high factor loadings on a major core depressive factor and adequate item discrimination values, supporting construct validity for not only the total sample, but also gender and age subgroups. Factor analysis of PHQ-9 with additional items revealed similarly acceptable loadings (0.43 to 0.63) among educated Nigerian army personnel[80].

Focus group interviews supported content validity, but suggested two potential modifications to the PHQ-9 instructions for western Kenyan culture (see Appendix). A common cultural practice in western Kenya is to respond "Nzuri" (fine) when asked how one is doing, even when one is not well. We especially wondered whether suicidal ideation would be self-reported. However, the mean for Item 9 was generally similar to US samples, relative to other item means (Fig. 2). Furthermore, focus groups appeared willing and relieved to share feelings.

Coefficient alpha of 0.75 to 0.81 for gender/age subgroups implies a high degree of internal consistency. Values of 0.80 and 0.85 were reported among Nairobian low-literacy adult cancer patients[11] and Nigerian university students[46], respec-

Table 2. Factor Analysis of PHQ-9: One Factor Solution

Item no.	Item name	Factor loading	Item-total Correlation	Item mean	Item SD
Part A. Males and females combined (n=347)					
1	Anhedonia	.65	.51	0.89	1.00
2	Depressed	.63	.49	0.78	0.98
3	Sleep problems	.58	.45	0.82	1.06
4	Low energy	.66	.52	1.10	1.00
5	Appetite problems	.57	.43	0.90	1.06
6	Low self esteem	.61	.47	0.89	1.13
7	Trouble concentrating	.64	.51	0.79	1.10
8	Psychomotor problems	.59	.45	0.64	0.97
9	Suicide ideation	.52	.39	0.38	0.81
Part B. Females only (n=251)					
1	Anhedonia	.63	.48	0.86	0.97
2	Depressed	.62	.47	0.80	0.98
3	Sleep problems	.55	.41	0.82	1.08
4	Low energy	.70	.55	1.15	0.99
5	Appetite problems	.51	.38	0.93	1.07
6	Low self esteem	.60	.45	0.94	1.15
7	Trouble concentrating	.65	.51	0.80	1.12
8	Psychomotor problems	.55	.41	0.68	0.97
9	Suicide ideation	.51	.37	0.41	0.84
Part C. Males only (n=93)					
1	Anhedonia	.69	.57	0.95	1.06
2	Depressed	.65	.51	0.71	1.00
3	Sleep problems	.65	.53	0.82	0.99
4	Low energy	.56	.44	0.95	1.00
5	Appetite problems	.67	.56	0.81	1.01
6	Low self esteem	.67	.55	0.73	1.03
7	Trouble concentrating	.62	.49	0.73	1.05
8	Psychomotor problems	.67	.55	0.53	0.95
9	Suicide ideation	.50	.38	0.27	0.66

The item-total correlation is corrected by excluding the item from the total score.

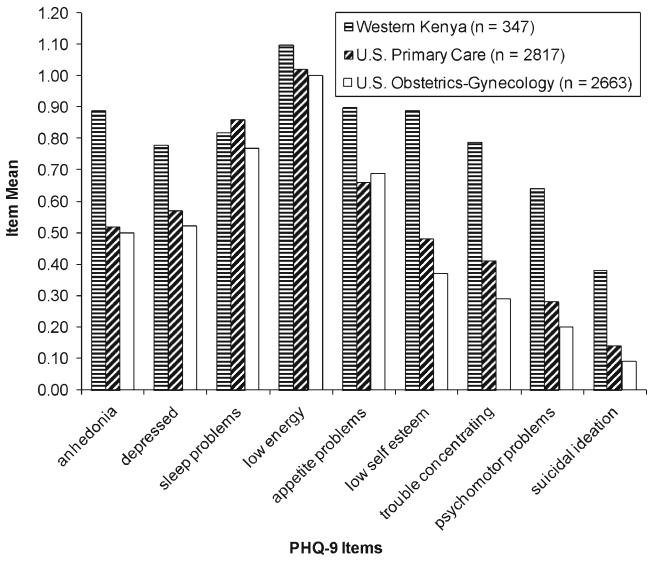

Figure 2. Mean score for each PHQ-9 item.

tively. Test-retest reliability of 0.59 for PHQ-9 total score is moderate. This ICC was better for older (0.65) and worse for younger (0.51) participants, and lower than previous PHQ-9 studies: 0.71 for a 14-day period among urban Kenyans with cancer[11], 0.89 for a 30-day period among Nigerian university students[46] and 0.76 to 0.92 for a 7-day period among US and European outpatient samples[56,57,62,73]. However, 0.59 may be acceptable given events between baseline and retest surveys that could have contributed to slight changes in depressive symptoms, including: (1) a national election with civil unrest and (2) some participants attended a support group session (data on who attended was not available).

Kappa coefficients of 0.21–0.40, 0.41–0.60 and 0.61–0.80 represent fair, moderate and substantial agreement, respectively[99]. Thus, MDD, ODD and any depressive disorder showed fair agreement. In the only study that reported test-retest reliability for PHQ-9 MDD, 7-day kappa was also not substantial (0.46)[62]. Weighted kappa (0.53) suggests there is moderate agreement for PHQ-9 severity categories among western Kenyans.

Consistent with previous research[58,76–78], the PHQ-2 demonstrated very good operating characteristics. The best cutoff point for sensitivity and specificity (≥3) provided very high specificity for any depressive disorder, an advantage in primary care and resource-constrained HIV clinics, where low specificity creates a problem by generating many false positives.

This study has several limitations. First, we did not assess gold standard diagnoses using a structured interview. Second, results can only be generalized to western Kenyan adults living with HIV/AIDS and, strictly speaking, to those attending HIV-related psychosocial support groups. Participant characteristics were similar to the AMPATH population from which participants were recruited. However, patients who attend support groups might have a better understanding of psychological symptoms and terminology. Third, language preference, variability in literacy and occasional assistance from research assistants may have impacted psychometrics and reproducibility of results. Moderate sample size prevented subgroup analyses for PHQ-2 operating characteristics and test-retest reliability of PHQ-9 diagnoses.

Future studies in western Kenya should assess PHQ-9 and PHQ-2 operating characteristics compared to structured diagnostic interviews and further psychometrics within demographic groups. The only study examining operating characteristics of PHQ-9 in sub-Saharan Africa found excellent results among Nigerian university students[46]. Studies should assess PHQ-9 responsiveness to treatment with anti-

Table 3. Operating Characteristics of PHQ-2 (n=347)

PHQ-2 cutoff point	TN	FN	FP	TP	Sensitivity	Specificity	Positive predictive value	Likelihood ratio
	PHQ-9 DSM-IV major depressive disorder (MDD)							
≥1	116	0	186	45	100.0	38.4	19.5	1.62
≥2	185	0	117	45	100.0	61.3	27.8	2.58
≥3	232	4	70	41	91.1	76.8	36.9	3.93
≥4	282	19	20	26	57.8	93.4	56.5	8.72
≥5	298	28	4	17	37.8	98.7	81.0	28.52
≥6	302	37	0	8	17.8	100.0	100.0	53.69
	AUC (95% CI)=0.91 (0.88–0.95)							
	Any PHQ-9 DSM-IV depressive disorder (MDD or ODD)							
≥1	116	0	113	118	100.0	50.7	51.1	2.03
≥2	185	0	44	118	100.0	80.8	72.8	5.20
≥3	218	18	11	100	84.7	95.2	90.1	17.64
≥4	228	73	1	45	38.1	99.6	97.8	87.33
≥5	229	97	0	21	17.8	100.0	100.0	40.75
≥6	229	110	0	8	6.8	100.0	100.0	15.53
	AUC (95% CI)=0.97 (0.95–0.98)							

ODD = Other depressive disorder. AUC = Area under receiver-operator curve.
TN = True negative = test negative (PHQ-2 < cutoff point), PHQ-9 disorder is absent.
FN = False negative = test negative (PHQ-2 < cutoff point), PHQ-9 disorder present.
FP = False positive = test positive (PHQ-2 ≥ cutoff point), PHQ-9 disorder absent.
TP = True positive = test positive (PHQ-2 ≥ cutoff point), PHQ-9 disorder present.
The likelihood ratio is undefined when specificity=100% and therefore was estimated for MMD at cutoff point 6 by assuming TN=301 and FP=1, and for any depressive disorder at cutoff points 5 and 6 by assuming TN=228 and FP=1.
By definition, PHQ-2 ≥1 and ≥2 produce FN=0 and sensitivity=100% when compared to PHQ-9 DSM-IV algorithm for MDD or any depressive disorder.

retroviral and psychosocial therapy, and whether PHQ-9 predicts adherence to antiretroviral therapy.

There are few mental health providers in Kenya[17]. Nevertheless, depression can be effectively treated in low-income countries[39,100,101]. The PHQ-9 appears to be a valid and reliable brief tool for assessing DSM-IV depressive disorders and depression severity in patients living with HIV/AIDS in resource-constrained western Kenya. Moreover, the PHQ-2 has good diagnostic operating characteristics as a two-item screener. The PHQ-9 and PHQ-2 may be useful to HIV-related medical providers in western Kenya who seek assessment tools for linking patients to psychosocial and psychiatric services in the continuum of care.

Acknowledgements: *This project was sponsored in part by funding provided by the Presidential Emergency Plan for AIDS Relief to the USAID-AMPATH® partnership. AMPATH is a registered trademark of the Trustees of Indiana University, Moi University and Moi Teaching and Referral Hospital. This project was also supported by funding from the School of Health, Physical Education, and Recreation and the Department of Applied Health Science at Indiana University-Bloomington.*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Patrick O. Monahan, PhD; Department of Medicine, Indiana University School of Medicine, 410 West 10th Street, Suite 3000, Indianapolis, IN 46202-3002, USA (e-mail: pmonahan@iupui.edu).*

REFERENCES

1. UNAIDS. Report on the global AIDS epidemic. Geneva: UNAIDS; 2006.
2. **Wild LG, Flisher AJ, Lombard C.** Suicidal ideation and attempts in adolescents: associations with depression and six domains of self-esteem. J Adolesc. 2004;27:611–24.
3. **Antelman G, Kaaya S, Wei R, et al.** Depressive symptoms increase risk of HIV disease progression and mortality among women in Tanzania. J Acquir Immune Defic Syndr. 2007;44:470–7.
4. **Mogga S, Prince M, Alem A, et al.** Outcome of major depression in Ethiopia: population-based study. Br J Psychiatry. 2006;189:241–6.
5. **Gureje O, Kola L, Afolabi E.** Epidemiology of major depressive disorder in elderly Nigerians in the Ibadan Study of Ageing: a community-based survey. Lancet. 2007;370:957–64.
6. **Bolton P, Neugebauer R, Ndogoni L.** Prevalence of depression in rural Rwanda based on symptom and functional criteria. J Nerv Ment Dis. 2002;190:631–7.
7. **Jelsma J, Mielke J, Powell G, De Weerdt W, De Cock P.** Disability in an urban black community in Zimbabwe. Disabil Rehabil. 2002;24:851–9.
8. **Hughes J, Jelsma J, Maclean E, Darder M, Tinise X.** The health-related quality of life of people living with HIV/AIDS. Disabil Rehabil. 2004;26:371–6.
9. **Jelsma J, Maart S, Eide A, Ka'Toni M, Loeb M.** The determinants of health-related quality of life in urban and rural isi-Xhosa-speaking people with disabilities. Int J Rehabil Res. 2007;30:119–26.
10. **Kaaya SF, Fawzi MCS, Mbwambo JK, Lee B, Msamanga GI, Fawzi W.** Validity of the Hopkins Symptom Checklist-25 amongst HIV-positive pregnant women in Tanzania. Acta Psychiatr Scand. 2002;106:9–19.
11. **Omoro SAO, Fann JR, Weymuller EA, Macharia IM, Yueh B.** Swahili translation and validation of the Patient Health Questionnaire-9 depression scale in the Kenyan head and neck cancer patient population. Int J Psychiatry Med. 2006;36:367–81.
12. **Ola BA, Adewuya AO, Ajayi OE, Akintomide AO, Oginni OO, Ologun YA.** Relationship between depression and quality of life in Nigerian outpatients with heart failure. Psychosom Res. 2006;61:797–800.
13. **Rosengren A, Hawken S, Ounpuu S, et al.** Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. Lancet. 2004;364:953–62.
14. **Smit J, Myer L, Middelkoop K, et al.** Mental health and sexual risk behaviours in a South African township: a community-based cross-sectional study. Public Health. 2006;120:534–42.
15. **Carson AJ, Sandler R, Owino FN, Matete FO, Johnstone EC.** Psychological morbidity and HIV in Kenya. Acta Psychiatr Scand. 1998;97:267–71.
16. **Sebit MB.** Neuropsychiatric HIV-1 infection study: in Kenya and Zaire cross-sectional phase I and II. Cent Afr J Med. 1995;41:315–22.
17. **Kiima DM, Njenga FG, Okonji MMO, Kigamwa PA.** Kenya mental health country profile. Int Rev Psychiatry. 2004;16:48–53.
18. **Olley BO, Seedat S, Nei DG, Stein DJ.** Predictors of major depression in recently diagnosed patients with HIV/AIDS in South Africa. AIDS Patient Care STDS. 2004;18:481–7.
19. **Olley BO, Seedat S, Stein DJ.** Persistence of psychiatric disorders in a cohort of HIV/AIDS patients in South Africa: a 6-month follow-up study. Psychosom Res. 2006;61:479–84.
20. **Els C, Boshoff W, Scott C, Strydom W, Joubert G, van der Ryst E.** Psychiatric co-morbidity in South African HIV/AIDS patients. S Afr Med J. 1999;89:992–5.
21. **Kaharuza FM, Bunnell R, Moss S, et al.** Depression and CD4 cell count among persons with HIV infection in Uganda. AIDS Behav. 2006;10(Suppl. 4):S105–S11.
22. **Keogh P, Allen S, Almedal C, Temahagili B.** The social impact of HIV infection on women in Kigali, Rwanda: a prospective study. Soc Sci Med. 1994;38:1047–53.
23. **Myer L, Smit J, Roux LL, Parker S, Stein DJ, Seedat 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;22:147–58.
24. **Olley BO, Gxamza F, Seedat S, et al.** Psychopathology and coping in recently diagnosed HIV/AIDS patients: The role of gender. S Afr Med J. 2003;93:928–31.
25. **Poupard M, Gueye NFN, Thiam D, et al.** Quality of life and depression among HIV-infected patients receiving efavirenz- or protease inhibitor-based therapy in Senegal. HIV Med. 2007;8:92–5.
26. **Ickovics JR, Hamburger ME, Vlahov D, et al.** Mortality, CD4 cell count decline, and depressive symptoms among HIV-seropositive women: longitudinal analysis from the HIV Epidemiology Research Study. JAMA. 2001;285:1466–74.
27. **Justice AC, McGinnis KA, Atkinson JH, et al.** Psychiatric and neurocognitive disorders among HIV-positive and negative veterans in care: Veterans Aging Cohort Five-Site Study. AIDS. 2004;18 (Suppl. 1):S49–59.
28. **Perry S, Fishman B.** Depression and HIV: how does one affect the other? JAMA. 1993;270:2609–10.
29. **Stoskopf CH, Kim YK, Glover SH.** Dual diagnosis: HIV and mental illness, a population-based study. Community Ment Health J. 2001;37:469–79.
30. **Reece M, Shacham E, Monahan P, et al.** Psychological distress symptoms of individuals seeking HIV-related psychosocial support in western Kenya. AIDS Care. 2007;19:1194–200.
31. **Shacham E, Reece M, Monahan P, Yebei V, Omollo O, Ong'or WO, Ojwang C.** Measuring psychological distress symptoms in patients living with HIV in Western Kenya. J Ment Health. in press.
32. **Shacham E, Reece M, Ong'or WO, Omollo O, Monahan P, Ojwang C.** Characteristics of psychosocial support seeking during HIV-related treatment in western Kenya. AIDS Patient Care STDS. 2008;22: 595–601.
33. World Health Organization. The world health report 2000. Mental health: New understanding, new hope. Geneva: World Health Organization; 2001.
34. **Freeman M.** HIV/AIDS in developing countries: Heading towards a mental health and consequent social disaster? S Afr J Psychol. 2004;34:139–59.
35. **Collins PY, Holman AR, Freeman MC, Patel V.** What is the relevance of mental health to HIV/AIDS care and treatment programs in developing countries? A systematic review. AIDS. 2006;20:1571–82.
36. **Bass JK, Bolton PA, Murray LK.** Do not forget culture when studying mental health. Lancet. 2007;370:918–9.
37. **Tomlinson M, Swartz L, Kruger L-M, Gureje O.** Manifestations of affective disturbance in sub-Saharan Africa: key themes. J Affect Disord. 2007;102:191–8.

38. **Uwakwe R, Okonkwo JEN.** Affective (depressive) morbidity in puerperal Nigerian women: validation of the Edinburgh Postnatal Depression Scale. Acta Psychiatr Scand. 2003;107:251–9.
39. **Verdeli H, Clougherty K, Bolton P, et al.** Adapting group interpersonal psychotherapy for a developing country: experience in rural Uganda. World Psychiatry. 2003;2:114–20.
40. **Wilk CM, Bolton P.** Local perceptions of the mental health effects of the Uganda acquired immunodeficiency syndrome epidemic. J Nerv Ment Dis. 2002;190:394–7.
41. **Rahim SIA, Cederblad M.** Epidemiology of mental disorders in young adults of a newly urbanized area in Khartoum, Sudan. Br J Psychiatry. 1989;155:44–7.
42. **Bolton P, Wilk CM, Ndogoni L.** Assessment of depression prevalence in rural Uganda using symptom and function criteria. Soc Psychiatry Psychiatr Epidemiol. 2004;39:442–7.
43. **Kagee A.** Symptoms of depression and anxiety among a sample of South African patients living with a chronic illness. J Health Psychol. 2008;13:547–55.
44. **Hamad R, Fernald LCH, Karlan DS, Zinman J.** Social and economic correlates of depressive symptoms and perceived stress in South African adults. J Epidemiol Community Health. 2008;62:538–44.
45. **Rochat TJ, Richter LM, Doll HA, Buthelezi NP, Tomkins A, Stein A.** Depression among pregnant rural South African women undergoing HIV testing. JAMA. 2006;295:1376–78.
46. **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.
47. **Adewuya AO, Ola BA, Dada AO, Fasoto OO.** Validation of the Edinburgh Postnatal Depression Scale as a screening tool for depression in late pregnancy among Nigerian women. J Psychosom Obstet Gynaecol. 2006;27:267–72.
48. **Pretorius TB.** Cross-cultural application of the Center for Epidemiological Studies Depression Scale: A study of Black South African students. Psychol Rep. 1991;69:1179–85.
49. **Awaritefe A.** The Beck Depression Inventory in relation to some commonly used tests in Nigeria. Niger J Basic Appl Psychol. 1988;1:23–8.
50. **Jelsma J, Mkoka S, Amosun L, Nieuwveldt J.** The reliability and validity of the Xhosa version of the EQ-5D. Disabil Rehabil. 2004;26:103–8.
51. **Rashid E, Kebede D, Alem A.** Evaluation of an Amharic version of the Composite International Diagnostic Interview (CIDI) in Ethiopia. Ethiop J Health Dev. 1996;10:69–77.
52. **Patel V, Todd C.** The validity of the Shona version of the Self Report Questionnaire and the development of the SRQ-8. Int J Methods Psychiatr Res. 1996;6:153–60.
53. **Spitzer RL, Kroenke K, Williams JBW, the Patient Health Questionnaire Primary Care Study Group.** Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. JAMA. 1999;282:1737–44.
54. **Spitzer RL, Williams JBW, Kroenke K, Hornyak R, McMurray J, the Patient Health Questionnaire Obstetrics-Gynecology Study Group.** Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. Am J Obstet Gynecol. 2000;183:759–69.
55. **Lee PW, Schulberg HC, Raue PJ, Kroenke K.** Concordance between the PHQ-9 and the HSCL-20 in depressed primary care patients. J Affect Disord. 2007;99:139–45.
56. **Pinto-Meza A, Serrano-Blanco A, Penarrubia MT, Blanco E, Haro JM.** Assessing depression in primary care with the PHQ-9: Can it be carried out over the telephone? J Gen Intern Med. 2005;20:738–42.
57. **Kroenke K, Spitzer RL, Williams JBW.** The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16:606–13.
58. **Williams LS, Brizendine EJ, Plue L, et al.** Performance of the PHQ-9 as a screening tool for depression after stroke. Stroke. 2005;36:635–38.
59. **Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL.** Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. J Gen Intern Med. 2006;21:547–52.
60. **Huang FY, Chung H, Kroenke K, Spitzer RL.** Racial and ethnic differences in the relationship between depression severity and functional status. Psychiatr Serv. 2006;57:498–503.
61. **Chen TM, Huang FY, Chang C, Chung H.** Using the PHQ-9 for depression screening and treatment monitoring for Chinese Americans in primary care. Psychiatr Serv. 2006;57:976–81.
62. **Fann JR, Bombardier CH, Dikmen S, et al.** Validity of the Patient Health Questionnaire-9 in assessing depression following traumatic brain injury. J Head Trauma Rehabil. 2005;20:501–11.
63. **Bombardier CH, Richards JS, Krause JS, Tulsky D, Tate DG.** Symptoms of major depression in people with spinal cord injury: Implications for screening. Arch Phys Med Rehabil. 2004;85:1749–56.
64. **Diez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL.** Validation and utility of the Patient Health Questionnaire in diagnosing mental disorders in 1,003 general hospital Spanish inpatients. Psychosom Med. 2001;63:679–86.
65. **Lowe B, Spitzer RL, Grafe K, et al.** Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. J Affect Disord. 2004;78:131–40.
66. **Martin A, Rief W, Klaiberg A, Braehler E.** Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. Gen Hosp Psychiatry. 2006;28:71–7.
67. **Wulsin L, Somoza E, Heck J.** The feasibility of using the Spanish PHQ-9 to screen for depression in primary care in Honduras. Prim Care Companion J Clin Psychiatry. 2002;4:191–5.
68. **Becker S, Al Zaid K, Al Faris E.** Screening for somatization and depression in Saudi Arabia: A validation study of the PHQ in primary care. Int J Psychiatry Med. 2002;32:271–83.
69. **Lowe B, Kroenke K, Herzog W, Grafe K.** Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). J Affect Disord. 2004;81:61–6.
70. **Dietrich AJ, Oxman TE, Burns MR, Winchell CW, Chin T.** Application of a depression management office system in community practice: a demonstration. J Am Board Fam Pract. 2003;16:107–14.
71. **Greco T, Eckert G, Kroenke K.** The outcome of physical symptoms with treatment of depression. J Gen Intern Med. 2004;19:813–8.
72. **Lowe B, Schenkel I, Carney-Doebbeling C, Gobel C.** Responsiveness of the PHQ-9 to Psychopharmacological Depression Treatment. Psychosomatics. 2006;47:62–7.
73. **Lowe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K.** Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care. 2004;42:1194–201.
74. **Rief W, Nanke A, Klaiberg A, Braehler E.** Base rates for panic and depression according to the Brief Patient Health Questionnaire: A population-based study. J Affect Disord. 2004;82:271–6.
75. **Glasgow RE, Nutting PA, King DK, et al.** A practical randomized trial to improve diabetes care. J Gen Intern Med. 2004;19:1167–74.
76. **Kroenke K, Spitzer RL, Williams JBW.** The Patient Health Questionnaire-2: Validity of a two-item depression screener. Med Care. 2003;41:1284–92.
77. **Lowe B, Kroenke K, Kerstin G.** Detecting and monitoring depression with a two-item questionnaire (PHQ-2). Psychosom Res. 2005;58:163–71.
78. **Corson K, Gerrity MS, Dobscha SK.** Screening for depression and suicidality in a VA primary care setting: 2 items are better than 1 item. Am J Manag Care. 2004;10:839–45.
79. **Katon WJ, Simon G, Russo J, et al.** Quality of depression care in a population-based sample of patients with diabetes and major depression. Med Care. 2004;42:1222–9.
80. **Okulate GT, Olayinka MO, Jones OBE.** Somatic symptoms in depression: evaluation of their diagnostic weight in an African setting. Br J Psychiatry. 2004;184:422–7.
81. **Siika AM, Rotich JK, Simiyu CJ, et al.** An electronic medical record system for ambulatory care of HIV-infected patients in Kenya. Int J Med Inform. 2005;74:345–55.
82. **Voelker R.** Conquering HIV and stigma in Kenya. JAMA. 2004;292:157–9.
83. **Wools-Kaloustian K, Kimaiyo S, Diero L, et al.** Viability and effectiveness of large-scale HIV treatment initiatives in sub-Saharan Africa: experience from western Kenya. AIDS. 2006;20:41–8.
84. **Mamlin J, Kimayo S, Nyandiko W, Tierney W.** Academic institutions linking access to treatment and prevention: case study. Geneva: World Health Organization. Retrieved August 26, 2008, from http://www.who.int/hiv/pub/prev_care/en/ampath.pdf; Nov 2004.
85. **Stewart AL, Hays RD, Ware JE.** The MOS short-form general health survey. Reliability and validity in a patient population. Med Care. 1988;26:724–35.
86. **Benjamini Y, Hochberg Y.** Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B Meth. 1995;57:289–300.
87. **Curran-Everett D.** Multiple comparisons: philosophies and illustrations. Am J Physiol Regul Integr Comp Physiol. 2000;279:R1–R8.

88. **Zwick R, Velicer WF.** Comparison of five rules for determining the number of components to retain. Psychol Bull. 1986;99:432–42.
89. **Howard KI, Forehand GA.** A method for correcting item-total correlations for the effect of relevant item inclusion. Educ Psychol Meas. 1962;22:731–5.
90. **Debus M.** Methodological review: a handbook for excellence in focus group research. Washington, DC: Academy for Educational Development; 1988.
91. **Willis GB.** Cognitive Interviewing: a tool for improving questionnaire design. Thousand Oaks, CA: Sage Publications, Inc.; 2005.
92. **McGraw KO, Wong SP.** Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1:30–46.
93. **Cohen J.** A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
94. **Cohen J.** Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968;70:213–20.
95. **Fleiss JL, Cohen J.** The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas. 1973;33:613–9.
96. **Dawson-Saunders B, Trapp RG.** Basic and clinical biostatistics. Norwalk, Connecticut: Appleton & Lange; 1990.
97. **Lyketsos CG, Hoover DR, Guccione M, et al.** Changes in depressive symptoms as AIDS develops: the Multicenter AIDS Cohort Study. Am J Psychiatry. 1996;153:1430–7.
98. **Nunnally JC, Bernstein IH.** Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.
99. **Landis JR, Koch GG.** The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
100. **Patel V, Araya R, Chatterjee S, et al.** Treatment and prevention of mental disorders in low-income and middle-income countries. Lancet. 2007;370:991–1005.
101. **Bass J, Neugebauer R, Clougherty KF, et al.** Group interpersonal psychotherapy for depression in rural Uganda: 6-month outcomes: randomised controlled trial. Br J Psychiatry. 2006;188:567–73.
102. **Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH.** The PHQ-8 as a measure of current depression in the general population. J Affect Disord. 2008:Aug 25 [Epub ahead of print, 1–11], doi:10.1016/j.jad.2008.06.026.

Appendix. PHQ-9 with Modified Instructions for Western Kenyan Culture

Based upon our focus groups, there are two minor modifications to the instructions that might be considered when administering the PHQ-9 in Western Kenya. First, some patients might benefit from being reminded that questions 3, 5, and 8 refer to changes in symptoms in *either direction*. Second, patients confused by the response options could be instructed that "not at all" refers to 0–1 days in the past 2 weeks, "several days" refers to 2–6 days, "more than half the days" refers to 7–11 days, and "nearly every day" refers to 12–14 days. This alternative *number of days* response set has been previously validated in a large population-based study of nearly 200,000 individuals.[102]

Over the last 2 weeks, how often have you been bothered by any of the following problems?

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things.....	0	1	2	3
2. Feeling down, depressed or hopeless.....	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much.....	0	1	2	3
4. Feeling tired or having little energy.....	0	1	2	3
5. Poor appetite or overeating.....	0	1	2	3
6. Feeling bad about yourself — or that you are a failure or have let yourself or your family down...	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television.....	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual.....	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way.....	0	1	2	3