

# Comparative Validity of Screening Instruments for Mental Distress in Zambia

Peter J. Chipimo<sup>1,2\*</sup> and Knut Fylkesnes<sup>2</sup>

<sup>1</sup>University of Zambia, School of Medicine, Department of Community Medicine, P/Bag RW X1, Lusaka, Zambia

<sup>2</sup>University of Bergen, Faculty of Medicine, Centre for International Health, 5020, Bergen, Norway

## Abstract:

**Background:** The recognition of mental health as a major contributor to the global burden of disease has led to an increase in the demand for the inclusion of mental health services in primary health care as well as in community-based health surveys in order to improve screening, diagnosis and treatment of mental distress. Many screening instruments are now available. However, the cultural validity of these instruments to detect mental distress has rarely been investigated in developing countries. In these countries, limited trained staff and specialized psychiatric facilities hamper improvement of mental health services. It is therefore imperative to develop a quick, low cost screening instrument that does not require specialized training. We validated different well established screening instruments among primary health care clinic attendees in Lusaka, Zambia. We also assess the face, content and criterion validity of the SRQ's and determined the most commonly reported symptoms for mental distress.

**Methods:** The screening instruments, SRQ-20, SRQ-10 and GHQ-12 were used as concurrent criteria for each other and compared against a gold standard, DSM-IV. Their correlation, sensitivity and specificity were assessed. All instruments were administered to 400 primary health care clinic attendees. In-depth interviews were also conducted with 28 of these clinic attendees.

**Results:** Both the SRQ-20 and SRQ-10 had high properties for identifying mental distress correctly with an AUC of 0.96 and 0.95 respectively while the GHQ-12 had modest properties (AUC, 0.81). The optimum cut-off points for this population were 7 and 3 for the SRQ and GHQ-12 respectively. The SRQ was also found to have good face and content validity.

**Conclusion:** The study establishes the utility of the SRQ-20 for detecting mental distress cases and also underscores the importance of validating instruments to suit the context of the target population. It also validates the SRQ-10 as the first reliable abbreviated and easy-to-use screening instrument for mental distress in primary health care facilities in Zambia.

**Keywords:** Mental distress, Screening instruments, Validity, Primary health care, SRQ-10, SRQ-20, GHQ12, DSM-IV, Zambia.

## BACKGROUND

Several investigations have shown that mental distress is common among health care seekers at primary health care centres but are not often identified, treated or referred [1]. Over the years, there has been increased attention to ways to improve the screening, diagnosis and treatment of mental distress in these patients. In many developing countries, trained staff and specialized psychiatric facilities are few and limited to urbanized areas [1]. Therefore in these countries, quick and low-cost means that do not require specialized training for assessing mental distress are essential. The ideal instrument should therefore be comprehensive, psychometrically sound and valid across cultures, age, sex, socio-economic and language background. This would require that the instrument be tested in different settings to enable comparisons between population groups within and across countries.

Among the most widely used self-administered tools are the Self Report Questionnaire (SRQ) and the General Health Questionnaire (GHQ) [2, 3]. Since the development of these instruments, detection rates for mental distress have steadily been increasing when employed in clinical settings or health surveys. Studies conducted in Ethiopia have revealed that between 6-18% of attendees at general outpatient clinics have mental distress [4-8]. These questionnaires have been tested in multicentre studies and have been translated into many languages [1, 3]. They have also been compared with other standardized psychiatric assessment in community based surveys and in primary care studies in developing countries [9, 10]. In Chile, the SRQ-20 and the GHQ-12 were simultaneously validated against the criterion of the Revised Interview Schedule (CIS-R) in a primary care setting. The results showed small differences between the SRQ and GHQ though the SRQ was found to be slightly more specific than the GHQ (77% vs. 73%) but closely comparable with regards to sensitivity (76% vs. 74%) [2]. A similar study in Brazil revealed the Pearson correlation between the two scales to be 0.72, with the validity coefficients for SRQ and GHQ being: sensitivity 83% vs. 85% and specificity 80% vs. 79% respectively. This study concluded that both

\*Address correspondence to this author at the University of Zambia, School of Medicine, Department of Community Medicine, P/Bag RW X1, Lusaka, Zambia; Tel: +47-96688704; Fax: +47-55 58 85 61; E-mail: Peter.Chipimo@cih.uib.no

instruments showed similar results [11]. The relatively few studies conducted in Sub-Saharan Africa have shown similar results, for example, Bhagwanjee *et al.* showed an unweighted sensitivity and specificity of 93.9% and 62.5% when the SRQ-20 was compared against the DSM-IV schedules for common mental disorders [12], while Reeler and Todd found sensitivity and specificity in the range of 80% [13]. Similar studies have been conducted among highly selected groups such as prenatal and postnatal women and in association with post-traumatic stress disorder in ex-combatants [14, 15]. From Zambia we could only find two studies which used the SRQ to measure mental distress. The first study validated the SRQ-20 by elucidating explanatory models for mental illness among low-income women while the other investigated the prevalence and determinants of mental distress and discussed the factors mediating its impact on HIV using the SRQ-10 as a screening instrument [14, 16]. Both studies, however, did not compare the SRQ to other established instruments and did not investigate the optimum cut-off point to be used for the Zambian population.

Most of these mental distress screening instruments started off as long, tedious and comprehensive scales which covered all dimensions of the universe of psychological/psychiatric constructs under study. However, with time they have been abbreviated in order to make them easy for use in busy clinic setting as well as in settings where some patients maybe illiterate and requiring the questionnaire to be read out to them. Emerging epidemiological studies investigating the correlation, reliability, the sensitivity and specificity between the long versions and the abbreviated versions of the instruments have shown that the later are just as capable (or even better) of identifying psychological distress. [17-19]. Good to excellent inter-rater agreement (Kappa coefficients) have been reported with abbreviated instruments and thus they have been judged to be acceptable and appropriate for use in different kinds of settings and countries [1, 20]. Overall these studies concluded that the subscales covering psychological distress functioned well and appeared to reflect a broad dimension of depression and anxiety disorders. The results also suggest that the shorter versions are valid and perform almost as well as the full versions, if not better, implying that these tools can be used inter-changeably, at least where depression is concerned [17, 18]. Along side considerations for an instrument's ability to identify cases, the factors that influence misclassification of cases also need due consideration. Several investigations have shown that misclassification by these questionnaires are significantly associated with social and demographic variables (education and sex), males being more likely than females to be misclassified as false negatives while the poorly educated respondents as false positives [2]. Other studies have attributed misclassification to language barriers, motives and cultural differences [21]. In a feasibility study conducted in Ethiopia using the SRQ-24, only moderate criterion validity was found. The limitations for this instrument in this study, was attributed partly due to it being very sensitive to help-seeking patterns of behavior by the participants. As a result, participants were found to be mentally distressed even in the absence of any mental illness. The study also revealed problems in trans-cultural communication because many of the diagnostic concepts used in this instrument were too "west-

ern" to be transposed unchanged to the Ethiopian culture. It was thus concluded that the items in the instruments needed fairly extensive modification to be applicable in the Ethiopian context [22].

In this paper we investigate the correlation, sensitivity and specificity, and we calculate the area under the curve (AUC) of receiver operating characteristics for various cut off points for the SRQ-20, SRQ-10 and GHQ-12 among primary health care clinic attendees in Lusaka, Zambia. The SRQ's and GHQ-12 are used as concurrent criteria for each other against the DSM-IV as the gold standard. We also assess the face, content and criterion validity of the SRQ's and determine the most commonly reported symptoms for mental distress in these scales.

## METHODS

### The Setting and Study Design

A concurrent nested mixed method research design was used (Fig. (1)). We assessed attendees at 4 primary health care centers run by the government of the republic of Zambia between December 2008 and May 2009. These clinics were purposely selected within the city of Lusaka, two of which were clinics in very high density areas (Kalingalinga and Mtendere) while the others were clinics in a medium density area (Chilenje and Chelston). The residents of these areas speak a number of languages but mainly English and Nyanja.

### Procedure

A pilot study was first conducted at Kabwata clinic (outside the study sites) (Fig. (1)). Forty-five outpatients were interviewed and based on the results it was decided that the questionnaire would be read to all the participants irrespective of their education level. A time sample of 400 clinic attendees aged 16 years and over was asked to participate in the study between January and March 2009. The purpose of the study was explained to each participant by the research assistants and consent was asked for. Each clinic was sampled randomly on selected days, 3 times each week. On the selected day, interviews were conducted with consecutive clinic attendees at the clinic outpatients department.

### Quantitative Procedures

A brief social and demographic questionnaire was administered to all the participants by research assistants who had received training in carrying out interviews. The interviews lasted approximately 10 minutes. Information on participant's demographic characteristics, including age, gender, educational attainment, residence and marital status, was collected using standard questionnaire items. The participants were also asked in what language they wanted the interview to be carried out. Socioeconomic position was assessed using the participant's educational attainment, employment status and an asset index based on items intended to reflect household wealth. These included household ownership of appliances (TV, radio, refrigerator, electricity, bicycle, plough, cattle and donkey) and other household resources (running water in the home, type of toilet, type of floor, and type of roofing material). A summative wealth index was then constructed which was categorized into low, medium and high wealth index. The participants were also asked to rate their own health status by answering the ques-

tion: How would you say your health is at the moment? Is it, (1) Very poor, (2) Poor, (3) Fair, (4) Good, or (5) Excellent? The recent life events were evaluated by events occurring in the previous 12 months based on whether the participant had experienced (1) Break-up of a marriage (2) Break-up of a sexual relationship, (3) Physical abuse, (5) Neglected or disowned by family or (6) loss of a loved one.

The SRQ-20 and the GHQ-12 were used to measure global mental distress. These interviews were conducted by interviewers of the same sex as the participant. The participants were then classified into two groups according to their scores on the SRQ-20 (low, 0-7; high 8+) and GHQ-12 (low, 0-3; high, 4+). Subsequently these participants were directed to a medical officer who held a clinical interview with them for the ailments that brought them to the clinic as well as conducting a psychiatric inquiry where the DSM-IV schedules for common mental disorders was used to determine the presence and diagnosis of a psychiatric disorder. The general health assessment and the mental distress assessment were done at the same time so that the patients were not delayed due to the study. The clinical interview was conducted blind, without the knowledge of the questionnaire results.

### Qualitative Procedures

In the second part of the study, in-depth interviews were conducted in a subsample of 28 participants nested within the quantitative sample. The sample consisted of participants who were classified as being high scorers (14 participants) and low scorers (14 participants), on the basis of the SRQ-20 score >7 and GHQ-12 score >3. These interviews were used to assess face and content validity.

### Face Validity

This facet simply indicates if on the face of it, the SRQ appears to assess meaningful and relevant qualities. Normally this facet is based on a review by a panel of experts. In its original development the SRQ was assessed by a panel of experts from different countries who selected SRQ items from different questionnaires. In this study the approach to assess face validity was to ask the target population what they think the instrument is suppose to measure.

### Content Validity

This consists of a determination of whether the instrument captures all the relevant concepts and if it is representative of the battery of questions that could have been asked for individuals under study. It is closely related to face validity since it also requires validation-by-assumption by a panel of experts. However the concept of content validity that we adopt here is a subjective judgment based on a review of the various items by the respondents. We thus asked the respondents to interpret their "yes" responses to the items in the SRQ-20. We also asked them to give us as many examples as possible to support their answer. We additionally asked them what remedy they think would work to abate the symptoms. Answers to these probing questions were used as a basis to ascertain whether the yes-answer had the same meaning for the respondent as it did for the investigator. The three stages considered in this study at which a yes-answer maybe invalid were the language of the interview, concepts

and motives behind the "yes" answer. The interviews took approximately 20 minutes per session.

## Instruments

### *Self-Reporting Questionnaire- 20 (SRQ-20)*

The SRQ-20 was developed by the World Health Organisation (WHO) as a screening tool for common mental disorders [1]. It was primarily developed for use in primary health care settings, especially in developing countries. Originally (SRQ-25) it consisted of 25 questions, 20 related to neurotic symptoms, 4 concerning psychosis and 1 asking about convulsions. This study concentrates on the SRQ-20, which (consists of 20 yes/no questions) assesses presence of neurotic symptoms (anxiety, depression, psychosomatic) mainly because few patients with functional psychosis come spontaneously to primary health centres and so usually more active case finding by primary health workers in the community is required. Secondly, psychotic patients are often easily recognised as being psychotic and in most cases, are unaware of their condition. Hence, the use of a questionnaire to detect psychoses is questionable. The SRQ-20 has been tested in numerous settings. Depending on the setting, community surveys or primary care, varied cut-off points have been used although cut-off point of 7/8 is widely used [1]. As far as we know no such study with equal representation of men and women has been conducted in Zambia.

### *Self-Reporting Questionnaire-10 (SRQ-10)*

The SRQ-10 is basically an abbreviated version of the SRQ-20. The instrument contains a weighted sum of 10 symptom questions which have dichotomous responses but do not probe to evaluate symptom severity. The scale measures the following symptoms over the preceding 30 days by asking: In the past 30 days: *Do you sleep badly?*, *Do you cry more than usual?*, *Do you find it difficult to enjoy your daily activities?*, *Do you find it difficult to make decisions?*, *Is your daily life suffering?*, *Are you unable to play a useful part in life?*, *Has the thought of ending your life been on your mind?*, *Do you feel tired all the time?*, *Do you often have headaches?*, *Is your digestion poor?* We have previously used this instrument in population based studies in Zambia and yielded results that were comparable to those of studies done using the SRQ-20 [16]. However, to our knowledge comparisons between the abbreviated versions and the full versions of Self-Reporting Questionnaires have not been done in Zambia and we could not find similar studies done elsewhere.

### *General Health Questionnaire- 12 (GHQ-12)*

The General Health Questionnaire is a screening instrument designed for use in general practice but has been shown to be valid for use in community surveys as well [19]. It was originally a 60 item questionnaire but subsequently a number of abbreviated versions have been derived. Thus, there are the 30-, 28-, 20- and 12- item versions. All these versions have been subjected to many validity studies and the authors reported validity indices that suggest that these are widely acceptable tools for detecting psychiatric morbidity. The

instrument contains 12 symptom questions which are scored on a four-point likert scale ranging (0-1-2-3) from much-less-than-usual to much-more-than-usual. However, in the analysis this scale is often collapsed to a dichotomous scale (0-0-1-1). Depending on the setting, community surveys or primary care, varied cut-off points have been used although cut-off point of 3+ is widely accepted as indicative of psychiatric morbidity [23].

### Gold Standard

#### *Diagnostic and Statistical Manual of Mental Disorder 4<sup>th</sup> Edition (DSM-IV)*

The Diagnostic and Statistical Manual of Mental Disorders (DSM) is the standard classification of mental disorders used by mental health professionals. It is intended to be applicable for use across settings, inpatient-outpatient clinics, primary care, and with community populations. It has been used by clinicians and researchers of many different orientations such as psychiatrists, psychologists, social workers, occupational and rehabilitation therapists, and other health and mental health professionals. It is also a necessary tool for collecting and communicating accurate public health statistics. The DSM has a diagnostic classification, which is the list of the mental disorders that are officially part of the DSM system and making a DSM diagnosis consists of selecting those disorders from the classification that best reflect the signs and symptoms that are afflicting the individual being evaluated. For each disorder, a set of diagnostic criteria indicating what symptoms must be present (and for how long) in order to qualify for a diagnosis are provided [24]. The use of these diagnostic criteria has been shown to increase diagnostic reliability (i.e. likelihood that different users will assign the same diagnosis) [23]. The DSM-IV is widely accepted and used as the gold standard for psychiatric diagnosis in Zambia.

### Instrument Translation

All the instruments were translated into Nyanja and Bemba as these are the most predominantly spoken languages in Lusaka. The results from the pilot study also confirmed that participants who did not speak English opted to be interviewed in Nyanja or Bemba. These instruments were then back translated to English by bilingual translators from the linguistics department of the University of Zambia. Discrepancies that were found were discussed further by a group that included the principle investigator, translators and a medical doctor from the psychiatric hospital. This was to ensure face validity as well as conceptual meaning. Few final changes were made after the pilot study.

### Training of Study Staff

A team of three male and three female interviewers who had no experience in mental health care administered the SRQ-20 and the GHQ-12. They, however, all had previous experience administering questionnaires in other epidemiological studies. A three day training session was conducted in administering the instruments. This involved explanation and discussion of conceptual definitions of each item in the instruments and role playing. This was followed by a 1 day field test.

### Ethical Issues

The Research and Bioethics Committee of the University of Zambia and the Ministry of Health, through the Lusaka District Health Office approved this study. Permission to conduct the study was also further sought from the authorities in charge of the Primary Health Centres. The study was conducted in accordance with the guidelines of Good Clinical Practices in biomedical research.

### Statistical Analysis

The data was analysed using SPSS version 15. In this study, receiver operating characteristic (ROC) analysis was used to identify a cut-off point for the SRQ-10, SRQ-20 and GHQ-12 as defined with the DSM-IV as the gold standard. This plots sensitivity against 1-specificity for each possible cut-off point. The sensitivity and specificity here being the fraction of true positive cases and true negative cases correctly identified by the screening tools respectively. Each ROC is characterised by an area under the curve (AUC) which generally indicates the overall accuracy of the questionnaire over a range of cut-off points to distinguish between cases and non-cases. AUC ranges between 0.0 to 1.0 with 1.0 indicating perfect prediction and 0.5 indicating a prediction equal to chance. Hence we used the AUC to compare the screening tools over the total range of scores. We performed a factor analysis with varimax rotation to check for measurement equivalence. This refers to the equivalence of construct or theoretical validities across populations, which is a prerequisite for the comparison of prevalence rates or mean scores of the scales [25]. Independent t-tests were performed to compare the scales between sexes while the Pearson Chi-square was used to compare the psychiatric diagnosis in the same groups. We also calculated Pearson correlation coefficients to examine the relationship between the scales.

## RESULTS

### Socio Demographic Characteristics

The sample was composed of 400 respondents who completed the SRQ-20 and the GHQ-12 and were subsequently referred to the Medical Doctor for clinical interview using the DSM-IV. These respondents were visiting the four Primary Health Care (PHC) centres for various medical reasons. Ten patients were not included because they refused the clinical interview. There were, however, no significant differences between the total sample and the participants that refused the clinical interview in sex ratio, wealth status, marital status and educational attainment. The respondents who were ethnically from the Bemba speaking tribes accounted for 26% of the total study population, while 16% were Nyanja and only 12% were Tonga. However almost half of the respondents preferred English as the language for the interview, while the others preferred Nyanja and Bemba (38.8% and 8.5% respectively). The sample had 167 (41.8%) men and 233 (58.3%) women (Table 1). The male patients ranged in age between 16 and 67 years with a mean of 32 years (SD=11.1). Female patients ranged between 16 and 65 years with a mean of 29 years (SD=9.4). The majority of participants were married (64%). Most of the patients had more than 8 years of education (secondary 56% vs. tertiary

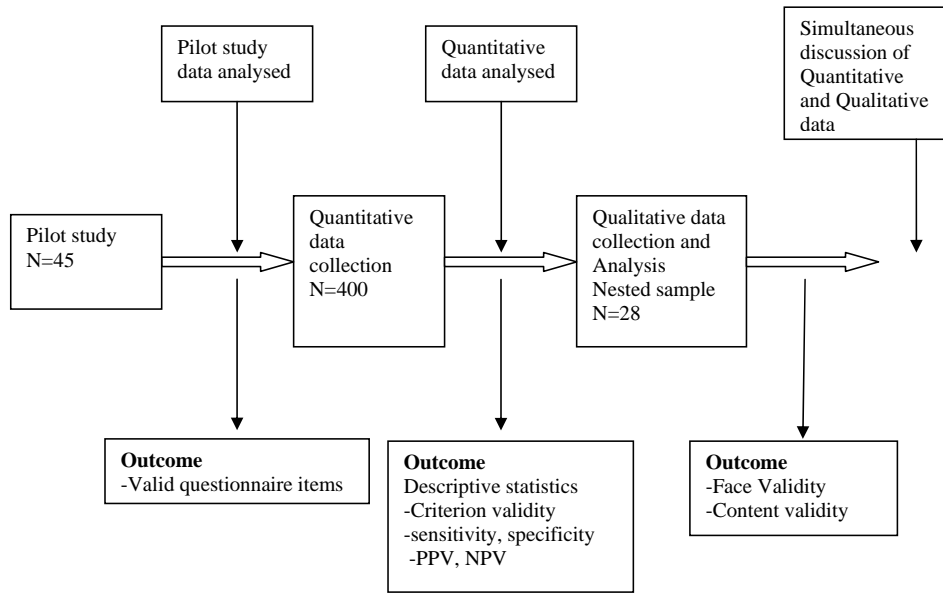


Fig. (1). Study design.

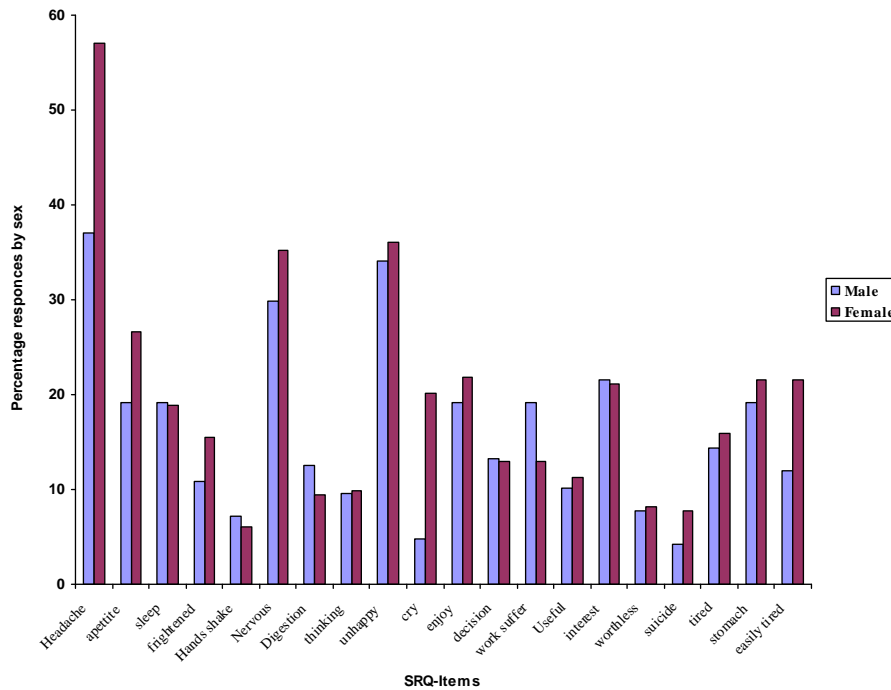


Fig. (2). Item by Item analysis of responses to SRQ-20.

19.5%) while 3.8% were illiterate. There were no statistical differences between the clinics serving the medium and high density catchment areas in terms of marital status ( $t= 1.139$ ,  $p=0.06$ ,  $\eta^2 =0.00$ ), wealth index ( $t=0.198$ ,  $p=0.418$ ,  $\eta^2 = 0.00$ ) and educational level ( $t=0.284$ ,  $p=0.777$ ,  $\eta^2 = 0.00$ ).

**Outcomes on SRQ-20, GHQ-12 and SRQ-10**

Principal component analysis with varimax rotation of the SRQ-20 items revealed a two factor model (common disorders and social disability) that explained 50.1% of the variance.

A similar model was extracted from the SRQ-10 and explained 50.2% of the variance, while three factors (Common

disorders, social dysfunction and loss of confidence) were extracted from GHQ-12 items by the same procedure explaining 49.9% of the variance. The factor structure of these instruments was similar to that reported in other studies [2, 26, 27]. We as a result found support for the measurement equivalence between the SRQ and GHQ-12 instruments. The correlation between the SRQ-20 and SRQ-10 was 0.85 while the correlation between these instruments and GHQ-12 scales was found to 0.60 and 0.52 respectively. Independent *t*-tests were used to compare differences in the continuous instrument scores between men and women and no significant differences were found. For comparison of definitive psychiatric diagnosis between males and females chi-square test was used and found to be insignificant ( $p=0.370$ ). Over-

Table 1. Social and Demographic Factors

Characteristic		Number (%) of Respondents		
		Male (N= 167)	Female (N= 233)	Total (N= 400)
<b>Age</b>	16-24	31.7	36.6	34.6
	25-29	13.2	25.9	20.6
	30-39	29.9	25.9	27.6
	40-49	16.2	7.3	11.0
	50+	9.0	4.3	6.3
<b>Marital status</b>	Single	44.3	30	36.0
	Married	55.7	70	64
<b>Education</b>	Illiterate	1.8	5.2	3.8
	Primary	11.4	27.2	20.6
	Secondary	60.5	53.0	56.1
	Tertiary	26.3	14.7	19.5
<b>Wealth index</b>	low	24.8	39.1	33.4
	middle	33.3	33.5	33.4
	High	41.8	27.4	33.2
<b>Language of Interview</b>	English	62.3	39.9	49.3
	Nyanja	29.3	45.5	38.8
	Bemba	6.0	10.3	8.5
	Other	2.4	4.3	3.5
<b>Gold standard</b>	DSM-IV	12.9	14.0	13.6
	Depression	11.0	11.0	11.0
	Anxiety	0.6	2.6	1.8

all the prevalence of common mental disorder as diagnosed by the DSM-IV classification was 13.6%, and was found to be mainly depression (10.8%) anxiety disorders (1.8%). The prevalence tended to be higher in females than males (women 14% vs. men 12.9%,  $p=0.743$ ). An item-by-item analysis of the SRQ also revealed that females on average reported more symptoms of mental distress than the males (Fig. 2).

### Criterion Validity

This part of the analysis focuses on the ability of the SRQ-20, GHQ-12 and SRQ-10 to screen for psychopathology (mental distress). Fig. (3) shows that SRQ-20 and SRQ-10 performed well with the area under the curve (AUC) being 0.96 and 0.95 respectively while the GHQ-12 had a modest AUC of 0.81. When analyzed separately for men and women no clear tendencies to perform better by sex were noted (Figs. 4, 5). Table 2 shows the sensitivity, specificity, positive predictive values, and negative predictive and kappa's values of the scales with different cut-off points. The most appropriate cut-off point was a trade off between sensi-

tivity and specificity. Since these instruments are meant to be used as screening instruments, the optimal cut-off point is one with high sensitivity and an acceptable specificity. The optimal cut-off for both SRQ-20 (sensitivity 0.85, specificity 0.94) and SRQ-10 (sensitivity 0.81, specificity 0.96) was 7, while that for GHQ-12 was 2 (sensitivity 0.66, specificity 0.86). Further analysis by sex did not reveal any significant differences in cut-off points.

### Content Validity of the SRQ

The study to assess the content validity was conducted in a subsample of the quantitative study. It included 28 respondents, 15 (53.6%) of whom were male while 13 (42.9%) were female. The respondents had an average of 9 school years being slightly higher in males than in females (10 years vs. 8 years respectively). Over half (53.6%) reported that they were married, 39.3% were single while less than 1% were either, divorced, separated or widowed. Half of the respondents preferred to have the interview conducted in English while 23% preferred Bemba while 28.6% preferred Nyanja. The 28 respondents gave the yes-answer a total of

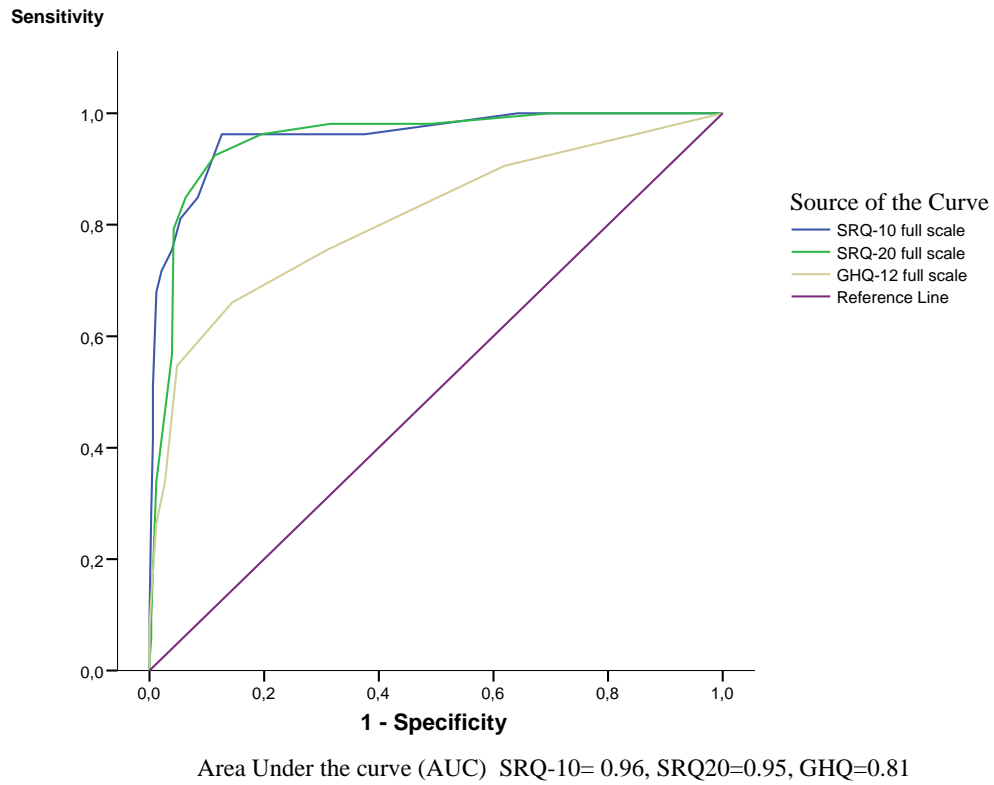


Fig. (3). Sensitivity, Specificity AUC for SRQ-20, SRQ-10 and GHQ-12 (combined).

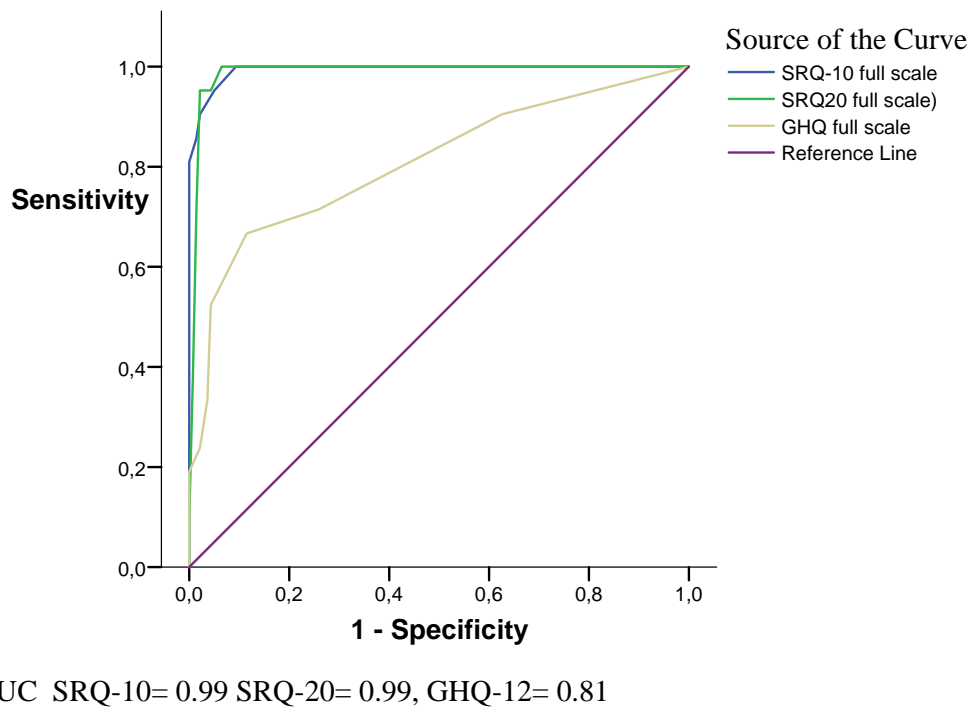
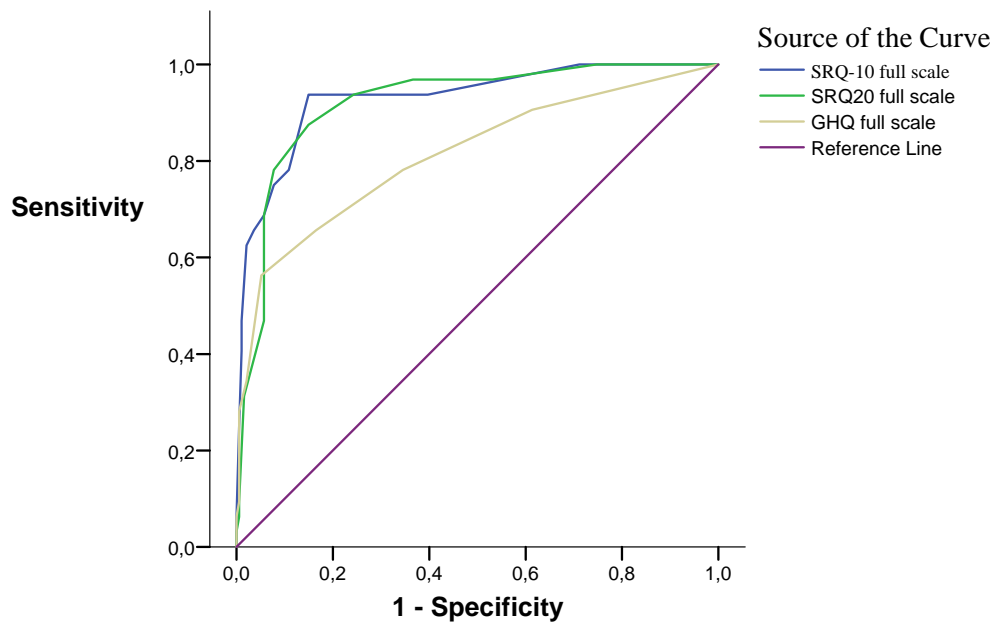


Fig. (4). Sensitivity, Specificity AUC for SRQ-20, SRQ-10 and GHQ-12 (Males).

205 times on the SRQ. Invalidity of these answers was considered on two main stages listed below. The results are presented in Table 3.

**Conceptualization**

Differences in conceptualization of the question by the respondent were recorded in 25% of the yes-answers given.



AUC SRQ-10=0.93, SRQ-20=0.92, GHQ-12= 0.81

Fig. (5). Sensitivity, Specificity AUC for SRQ-20, SRQ-10 and GHQ-12 (Females).

Table 2. Criterion Validity for SRQ-20, SRQ-10 and GHQ-12

	Scale	Cut-off	Sensitivity	Specificity	PPV	NPV	% of Cases Screened Correctly	k	%Cases
<b>Total</b>	SRQ-20	7	0.85	0.94	0.68	0.97	92.6	0.71	16.5
		8	0.79	0.96	0.75	0.97	93.6	0.73	14.0
		9	0.57	0.96	0.70	0.93	90.8	0.57	10.8
	SRQ-10	7	0.81	0.95	0.71	0.97	92.8	0.71	15.3
		8	0.76	0.96	0.76	0.96	93.3	0.72	13.3
		9	0.71	0.98	0.84	0.96	94.4	0.74	11.3
	GHQ-12	2	0.66	0.86	0.43	0.94	83.2	0.42	21.6
		3	0.57	0.95	0.67	0.93	90.2	0.56	11.9
		4	0.34	0.97	0.67	0.90	88.6	0.39	6.8

“Do you have headaches often?” All the invalid answers given to this question were attributed to the presence of other intercurrent illness namely hypertension, malaria and toothaches. However the question largely managed to uncover information indicating the headache as a symptom of depression and/or anxiety.

“Do you have uncomfortable feelings in your stomach?” - Among those giving invalid answers, this question was understood as an inquiry into presence of gastrointestinal ailment. The reasons most frequently given were: “Yes because I suffer from “gas” in my stomach” and “Yes I get uncomfortable feeling when I eat beans”. Contrary the questions; “Is your digestion poor?” and “Is your appetite poor?” performed very well with the most frequent answer among

the valid answers being: “Yes, I don’t feel like eating because I have many thoughts and even when I feel like eating I have problems swallowing or I get full easily”.

Anxiety Items: “Are you easily frightened? Do your hands shake? Do you feel tense or worried?” These items seemed to have a narrow meaning in the context of our study, and were interpreted as being an enquiry into literal feeling or state of being afraid. The most frequent answer was: “Sometimes, especially if I am threatened or if I am in trouble with spouse”. We also probed the no-answers to these items and we found the same responses suggestive of the fact that being frightened, hands shaking or feeling tense or worried is associated with literal fear. This concept does not seem to exist in our sample unless there is a clear reason



**Table 3. Content Validity of SRQ**

SRQ-Items	Yes-Answers N= 205	Reasons for Invalid Answers		
		Concepts n (%)	Language/Motives n (%)	Total (%)
1. Headache*	16	4 (25)	-	25
2.Appetite	9	2 (22.2)	2 (22.2)	44.4
3. Sleep*	17	3 (17.6)	1 (5.9)	23.5
4. Easily frightened	5	5 (100)	-	100
5. Hands shaking	5	3 (60)	2 (40)	100
6. Feel nervous	6	1 (16.7)	-	16.7
7. Poor Digestion*	7	2 (28.6)	5 (71.4)	100
8. Trouble thinking clearly	11	1 (9)	1 (9)	18.2
9. Unhappy	16	-	-	-
10. Cry more*	11	1 (9)	1 (9)	18.2
11. Enjoy activities*	7	-	2 (28.6)	28.6
12. Difficulty deciding*	5	-	1 (20)	20
13. Work suffering*	16	4 (26.7)	2 (13.3)	37.5
14. Useful in life*	13	1 (7.7)	4 (30.8)	38.5
15. Loss of interest	10	4 (40)	2 (20)	60
16. Worthlessness	10	-	-	-
17.Thoughts of suicide*	5	-	-	-
18. Always tired*	9	7 (77.8)	1 (11.1)	88.9
19. Stomach	8	4 (50)	3 (37.5)	87.5
20.Easily tired	9	7 (77.8)	1 (11.1)	88.9
Depression Items <sup>§</sup> (Items 10, 11, 12, 13, 14, 17)	57	6 (10.5)	10 (17.9)	28.4
Somatic items <sup>§</sup> (Items 1,3, 7, 18)	49	16 (32.6)	7 (14.3)	46.9

\* <sup>§</sup> Items included in SRQ-10.

for it and so the items failed to uncover the information suggestive of anxiety.

“Do you feel tired all the time?” Was interpreted by the respondents as asking about whether they get tired easily as regards work rather than an enquiry pertaining to depression. The most frequent answer was “Yes I get tired because of work since I work very long hours”.

### Language and Motives

We assigned a yes-answer to this invalidity category if the question had to be repeated one or more times or if it needed further explanation before an answer was obtained. We also assigned, to this category, respondents who said they didn't understand or who answered “I do not know” to the questions posed. We also included in this category respondents who insisted on the yes-answer but were unable or unwilling to give further details or examples of experiences

that would help us to clearly define the underlying psychopathology. Respondents who also directly indicated that they thought by participating in the interview they would be “fast-tracked” to see the doctor were also assigned to this category, although these accounted for less than 1%. This kind of invalid answers were observed in 15.6% of the yes-answers and was attributed to not understanding the content of the question and complexity of the words used.

### Face Validity of SRQ

Within the subsample we also assessed the face validity of the SRQ by asking the respondents what they thought the instrument was supposed to measure and we also probed further by asking the respondents what they thought the aim of these questions were. The SRQ was found to have good face validity with 71.4% of the respondents saying that we were assessing mental health. The most common response

was that we were measuring “problems of the mind and soul” (53.6%) while 17.9% said we were assessing stress and depression. The proportion who said they did not know the aim of the questions was 28.6%.

## DISCUSSION

We employed a concurrent nested mixed methods research design (QUAN qual) in a cross-sectional study conducted in four primary health care centers in the city of Lusaka aimed at comparing the validity of the SRQ-10 against that of the SRQ-20 and GHQ-12 in the screening for mental distress. DSM-IV was used as the gold standard. Overall the SRQ-10 showed good criterion validity at the optimum cut off point of 6/7 with the area under the curve (AUC) being 0.96 with good sensitivity and specificity (0.85 and 0.94 respectively). It was highly correlated to the SRQ-20 and only modestly to GHQ-12. (0.85 vs. 0.52) The SRQ-10 was also found to have good face validity. Content invalidity was found surrounding the anxiety items (Frightened, hands-shaking and nervous) and some somatic items (Headache, abdominal symptoms and tiredness). This was attributed mostly to conceptualization and to a less extent Language and motives. The prevalence of mental distress was found to be 13.6% compared with 15.3% based in the SRQ-10. This point prevalence is close to what was found in a population survey conducted in Zambia [16], and falls within the range of reported prevalence of mental distress in the region [28, 29].

We compared the abbreviated SRQ-10 with the widely validated SRQ-20. Different validation coefficients have been reported for the SRQ-20 in these countries [1, 11]. A study in Kenya validated SRQ against the Clinical Interview Schedule (CIS) and reported specificity of 93.3% and specificity of 89.2% [10], while a study in Ethiopia reported a sensitivity range of 68.4%-85.7% and specificity ranging between 62%-75.6% when they validated the SRQ against the Edinburgh postnatal depression scale (EPDS) [15]. In our study we found very high correlation coefficient between SRQ-10 and SRQ-20 with similarly high validation coefficients. The minor differences in the coefficients could be due to the use of different gold standards. It might also be attributed to the differing samples to which the instruments were applied. The validation coefficients reported here might also be somewhat higher because the study was conducted in an urban population with an average of education of 8 years and 50% of whom preferred English as the language of the interview. Comparison of the SRQ-10 and the GHQ-12 revealed a rather modest correlation coefficient despite GHQ-12 having acceptable validation coefficients. The validation coefficients we found were lower than those reported in other studies [2, 19]. This could be attributed to the negative phasing of its items. Often the questions had to be rephrased several times for the respondent to understand. The likert scale also proved to be confusing for the respondents and challenging to score for the research assistants. This challenge with scoring the GHQ-12 has also been reported by other authors who have questioned the best method of scoring [30-32] and the value of the using the likert scoring system [29, 32]. Another plausible reason is the cut off point we used for the GHQ-12. Although the cut off point we used is similar to that used in other studies, evidence suggests that using the median score

as the cut off point is better than using the mean score or other predetermined cut off points, especially in population which are “GHQ naïve” [33].

Broadly speaking, the validity coefficients did not seem to be affected by the socio-demographic factors as there were no statistically significant relationships noted. It was therefore unnecessary to use a different cut off point for men and women. These findings are different from some other studies that have suggested a higher false negative rate in men than in women attributed to the fact that expression of emotion would be stigmatizing among men [11].

The SRQ-10 showed good criterion validity overall although a limited percentage of participants gave invalid answers to some items on somatic symptoms. Several reasons can be given to explain this but the most important seemed to be related to communication problems based on different conceptual meaning. Improvement of the translation and further adjustment tailored to culturally understandable concepts may solve this problem. Other studies have reported poor criterion validity possibly related to health seeking behavior of the clinic attendees, i.e. a tendency to give more yes-answers in an attempt to receive special attention, a medical certificate or in order to be “fast-tracked” along in the queue [21, 22]. However, this was not revealed in our study. The anxiety items on the SRQ-20 appeared to have performed poorly, a finding that has also been reported in other studies as well. An investigation in Lesotho reported similar low reporting of anxiety symptoms due in part to poor understanding of the anxiety items. Respondents in this study tended to be moderately impaired by anxiety and often reported that they did not know what caused their symptoms [29]. It was suggested that understanding of these items can be enhanced by adjusting and translating the items into a locally palatable context. These anxiety items are however, not part of the SRQ-10 and the benefits of including them into the SRQ-10 were not immediately apparent. Literature has shown before that depressive disorders in Sub-Saharan Africa are more common than anxiety disorders [29, 34]. This has been confirmed in our study. It has also been reported that generalized anxiety disorders presents mainly as a mixed syndrome with depressive features in developing countries. A simple assumption can therefore be made that the depressive items in the SRQ-10 will also capture cases of anxiety disorder [1, 12, 35].

This study has limitations and strengths. Participants were restricted to urban settings with relatively high education attainment compared to rural populations. The external validity of the validation results might be difficult to judge. However, the instrument seems to be rather robust and the findings were closely related to studies conducted in a variety of communities, and this gives an indication that these findings can be extrapolated to the national level and even above – to the regional level. Furthermore, the sample size was relatively small and future validations should consider employing larger sample sizes. The main strength of the study stems from the fact we were able to draw upon universally acceptable etic instruments (SRQ-20 & GHQ-12) which have been used extensively in various countries and cultural orientations as comparatives for the SRQ-10. We also made an effort to strengthen the clinical and cultural validity via a

standard translation and back-translation process and ensuring retention of the original meaning of the questions. This process gave us reasonable confidence to use these instruments across cultures [36, 37]. We also adopted a concurrent nested mixed methods design which was a powerful tool in illuminating the content validity of the SRQ-items, hence supplementing the overall strength of these results. We believe therefore that these validation results can form a valid and reliable basis for further research in this field in the region.

## CONCLUSION

The present study has found that the SRQ-10 is a practical tool for measuring mental distress in primary health care. It has been shown to be robust when compared to other widely validated tools. (SRQ-20& GHQ-12) It has also been shown that the dichotomous response system appears to hold an advantage over the likert scales as it appeared to be easier to understand and yielded better results than those of an instrument scored on a likert scale. (GHQ-12) This has been shown to be true in other studies as well where the instruments were used for screening purposes [16, 29]. The SRQ-10 also holds an operational advantage as it is a shorter scale making it a more attractive option for use in busy primary health care services, in mental health surveys and also in general health surveys. To cover the whole range of mental disorders or to make diagnosis, it is imperative that it is coupled with other more comprehensive diagnostic scales [1].

## IMPLICATIONS OF THE STUDY

It has been reported previously that somatic symptoms associated with physical illness are often signs of mental distress [1, 37, 38]. In our study the respondents did not come to the clinic primarily for mental health problems but for other physical illnesses. This underscores the usefulness of screening questions for mental distress to patients with various medical conditions as this will help to identify at-risk individuals. The study is also a call for the adoption of the SRQ-10 as preferred simple, straightforward protocol screening tool as most mental health screening tools are long and tedious imposing unbearable strain on the busy and understaffed health workers. We feel that the question items can easily be incorporated into existing patient assessment protocols, thus enhancing case finding at primary health care level.

## LIST OF ABBREVIATIONS

SRQ-20	=	Self-Reporting Questionnaire 20
SRQ-10	=	Self-Reporting Questionnaire-10
GHQ-12	=	General Health Questionnaire
DSM-IV	=	Diagnostic and Statistical Manual of Mental Disorder 4 <sup>th</sup> Edition
QUAN	=	Quantitative
qual	=	qualitative
ROC	=	Receiver Operating Characteristics
AUC	=	Area Under the Curve

## ACKNOWLEDGEMENTS

The authors would like to acknowledge David Sam Lackland for advice and in the critical revision of the final draft.

We would also like to acknowledge the financial support from the Norwegian Programme for Development, Research and Education (NUFU).

## REFERENCES

- [1] World Health Organisation. A user's guide to the Self-Reporting Questionnaire (SRQ-20). World Health Organisation. Geneva 1994.
- [2] Araya R, Wynn R, Lewis G. Comparison of two self administered psychiatric questionnaires (GHQ-12 and SRQ-20) in primary care in Chile. *Soc Psychiatr Epidemiol* 1992; 27:168-73.
- [3] Goldberg D. The detection of psychiatric illness by questionnaire. London: Oxford University press 1972.
- [4] Giel R, Gezahegn Y, Van-Luuk JN. Psychiatric morbidity in 200 Ethiopian medical outpatients. *Psychiatr Neural Neurochir* 1968; 71: 169-76.
- [5] Giel R, Van-Luuk JN. Psychiatric morbidity in a small Ethiopian town. *Br J Psychiatry* 1969; 115:149-62.
- [6] Giel R, Van-Luuk JN. Pattern of Morbidity in general outpatient population as a guide to priorities in a medical curriculum. *Trop Geogr Med* 1968; 20: 281-7.
- [7] Dormaar M, Giel R, Van-Luuk JN. Psychiatric illness in two contrasting Ethiopian outpatient populations. *Soc Psychol* 1974; 9:155-61.
- [8] Kortmann F. Problems in practicing psychiatry in Ethiopia. *Ethiop Med J* 1988; 26: 77-83.
- [9] Sen B, Wilkinson G, Mari J. Psychiatric morbidity in primary health care: a two-stage screening procedure in developing countries. Cost effectiveness and choice of Instruments. *Br J Psychiatry* 1987; 151: 33-8.
- [10] Dhadphale R, Ellison R, Griffin L. The frequency of psychiatric disorders among patients attending semi-urban and rural general out-patients clinics in Kenya. *Br J Psychiatry* 1983; 142:379-83.
- [11] Mari J, Williams P. A validity study of a psychiatric screening questionnaire (SRQ-20) in Primary Care in the city of Sao Paulo. *Br J Psychiatry* 1986; 148: 23-6.
- [12] Bhagwanjee A, Parekh A, Paruk Z, Petersen I, Subedar H. Prevalence of minor Psychiatric disorder in an adult African rural community in South Africa. *Psychol. Med.* 1998; 28: 1137-47.
- [13] Reeler AP, Todd CH. An overview of Psychological disorders in psychiatric services in Zimbabwe. In proceedings of the primary mental health care workshop: University of Durban-Westville. Dept of Psychology: Kwazulu Natal 1994.
- [14] Aidoo M, Harpham T. The explanatory models of mental health amongst low income women and health care practitioners in Lusaka, Zambia. London: Oxford University press 2001; 16(2): 206-13.
- [15] Hanlon C, Girmay M, Atalay A, *et al.* Detecting perinatal common mental disorders in Ethiopia: validation of the self-reporting questionnaire and the edinburgh postnatal depression scale. *J Affect Disord* 2008; 108: 251-62.
- [16] Chipimo PJ, Fylkesnes K. Mental distress in the general population in Zambia: impact of HIV and social factors. *BMC Public Health* 2009; 9: 298.
- [17] Strand HS, Dalgard OS, Tambs K, Rognerud M. Measuring the mental health status of the Norwegian population: A comparison of instruments SCL-25, SCL-10 and MHI-5 (SF-36). *Nord J Psychiatry* 2003; 57(2): 113-8.
- [18] Tambs K, Moum T. How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatr Scand* 1993; 87(5): 364-367.
- [19] Goldberg DP, Gater R, Sartorius N, *et al.* The validity of two versions of GHQ in the WHO study of mental illness in general health care. *Psychol Med* 1997; 27: 191-7.
- [20] Tafari SAF. Determinants of mental illness in a rural Ethiopian population. *Soc Sci Med* 1991; 32 (2): 197-201.
- [21] Kortmann F, Ten-Horn S. Comprehension and motivation in responses to a psychiatric screening instrument. Validity of the SRQ in Ethiopia. *Br J Psychiatry* 1988; 153: 95-101.

- [22] Kortmann F. Psychiatric case finding in Ethiopia: shortcomings of the self reporting questionnaire. *Cult Med Psychiatry* 1990; 14 (3): 381-91.
- [23] Mario M, Wolfgang G, Juan-Jose L, Norman S. Psychiatric diagnosis and classification. 3rd ed. John Wiley & Sons 2002.
- [24] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). 4<sup>th</sup> ed. Washington DC 2000.
- [25] De Jong JTVM, Komproue IH, Spinazolla J. DESNOS in 3 postconflict settings: assessing cross-cultural construct equivalence. *J Trauma Stress* 2005; 18: 13-21.
- [26] Gao F, Luo N, Thumboo J, Fones C, Li S, Cheung Y. Does the 12-item general health questionnaires contain multiple factors and do we need them? *Health Qual Life Outcomes* 2004; 2(63): 1-7.
- [27] Politi P, Piccinelli M, Wilkinson G. Reliability, validity and factor structure of the 12 item General Health Questionnaire among young males in Italy. *Acta Psychiatr* 1994; 90(4): 432-7.
- [28] Giel R, van-Luuk J.N. Psychiatric morbidity in a small Ethiopian town. *Br J Psychiatry* 1969; 115: 149-62.
- [29] Hollifield M, Katon W, Spain D, Pule L. Anxiety and depression in a village in lesotho, africa: a comparison with the united states. *Br J Psychiatry* 1990; 156: 343-50.
- [30] Gureje O, Obikoya B. The GHQ as a screening tool in a primary care setting. *Soc Psychiatry Psychiatr Epidemiol* 1990; 25: 276-80.
- [31] Pan P, Goldberg DP. A comparison of the validity of GHQ-12 and CHQ-12 in Chinese primary care patients in Manchester. *Psychol Med* 1990; 20: 931-40.
- [32] Piccinelli M, Bisoffi G, Bon M, Cunico L, Tansella M. Validity and test-retest reliability of the Italian version of the 12-item General Health Questionnaire in general practice: a comparison between three scoring methods. *Comprehensive Psychiatry* 1993; 34: 198-205.
- [33] Willmott SA, Boardman JA, Henshaw CA, Jones PW. Understanding General Health Questionnaire (GHQ-28) score and its threshold. *Soc Psychiatry Psychiatr Epidemiol* 2004; 39(8): 613-7.
- [34] Mayeya J, Chazulwa R, Mayeya PN, *et al.* Zambia mental health country profile. *Int Rev Psychiatry* 2004; 16(1): 63-72.
- [35] Kleinman A, Good B. Culture and depression. Los Angeles: University of California Press 1985.
- [36] Ngoma MC, Prince M, Mann A. Common mental disorders among those attending primary health clinics and traditional healers in Urban Tanzania. *Br J Psychiatry* 2003; 183: 349-55.
- [37] Gureje O, Obikoya B. Somatisation in primary care: pattern and correlates in a clinic in Nigeria. *Acta Psychiatr* 1992; 86(3): 223-7.
- [38] Sayar K, Kirmayer LJ, Taillefer SS. Predictors of somatic symptoms in depressive disorder. *Gen Hosp Psychiatr* 2003; 25(2): 108-14.

Received: November 16, 2009

Revised: November 23, 2009

Accepted: November 23, 2009

© Chipimo and Fylkesnes; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.