

## Validation of the Sinhala version of 30-item General Health Questionnaire

Chrisantha Abeysena <sup>1</sup>, Upali Peiris <sup>2</sup>, Pushpa Jayawardana <sup>3</sup>, Asiri Rodrigo <sup>4</sup>

<sup>1</sup> Senior Lecturer / Department of Public Health, Faculty of Medicine, University of Kelaniya, Sri Lanka

<sup>2</sup> Senior Lecturer / Department of Psychiatry, Faculty of Medicine, University of Kelaniya, Sri Lanka

<sup>3</sup> Professor / Department of Public Health, Faculty of Medicine, University of Kelaniya, Sri Lanka

<sup>4</sup> Lecturer / Department of Psychiatry, Faculty of Medicine, University of Kelaniya, Sri Lanka

**Corresponding Author:** Dr. Chrisantha Abeysena

Tel: 2953411 (Department) | Mobile: 0777 591715

Email: chrisantha-abeyseana@mfac.kln.ac.lk, chrisanthaabeysena@yahoo.com

---

### Abstract

**Background:** The General Health Questionnaire (GHQ) is a widely used instrument to detect non potential psychotic psychological problems. The Sinhala translation of GHQ-30 has been widely used for research purposes in Sri Lanka, but details pertaining to the validation procedure have not been published.

**Objective:** To determine the criterion validity of Sinhala versions of GHQ-30.

**Methods:** This was a descriptive cross-sectional study including 374 patients aged 18 to 75 years, attending Out Patient Department of Colombo North Teaching Hospital between June 2009 to September 2010. Self administered Sinhala version of GHQ-30 was completed by the participants. Clinical Examination was done blindly to the GHQ score by a Psychiatrist based on ICD 10 classification of Depression/Anxiety/social dysfunction as reference standard. Cutoff values for GHQ-30 was determined by applying Receiver Operation Characteristic (ROC) curves.

**Results:** Optimal cutoff levels were 5/6 for GHQ-30 with a sensitivity of 67.5% (95%CI: 59%–76%) and a specificity of 80% (95%CI: 75%–85%). Optimal cutoff levels were not changed by sex, age and educational levels. The area under the ROC curve for GHQ-30 was 0.79. Despite above, multilevel likelihood ratios of >1 were obtained only for total scores of  $\geq 7$  for GHQ-30.

**Conclusions:** GHQ 30 displays adequate validity for use in the Sinhala speaking primary care setting attendees for assessing psychiatric disorders.

---

**Key words:** General Health Questionnaire, mental disorders, primary care, stress, validity

---

## **Introduction**

The General Health Questionnaire (GHQ) is a widely used instrument to detect non potential psychotic psychological problems.<sup>1</sup> The GHQ is only sensitive to assess short-term psychiatric disorders in the community or non-psychiatric clinical settings. The original version of this consisted of 60 question items, but several shortened versions have been derived from it. The main disadvantage of GHQ-60 is the time required to complete it, especially for those with poor reading skills as it is meant to be a self administered instrument. The GHQ-30 is an excellent quick screener and was developed removing all items related to somatic symptoms. This has been translated into Sinhala language and used widely in Sri Lanka for both research as well as for clinical purposes.

When used as a screening instrument it is imperative that they are valid and reliable. Validity is the extent to which a test measures what it claims to measure which is vital for a test to be valid in order for the results to be accurately applied and interpreted. A test is said to have criterion validity when the test is demonstrated to be effective in predicting criterion or indicators of a construct<sup>12</sup> by comparing it with a referenced standard. This indicates the extent to which the test scores accurately estimate an individual's current state with regard to the criterion.

The GHQ has been validated in different languages and cultures<sup>3,4</sup> and at different settings. However, the cutoff scores required to achieve the optimum sensitivity and specificity varies considerably from one setting to another.

Even though the Sinhala translation of GHQ-30 has been widely used for research purposes in Sri Lanka, details pertaining to the validation procedure have not been documented. Thus the objective of this study was to determine the criterion validity of the Sinhala version of the GHQ-30.

## **Methods**

A descriptive cross sectional study was conducted at the Out Patient Department (OPD) of the Colombo North Teaching Hospital (CNTH), Ragama between June 2009 to September 2010. The study participants included both males and females aged 18 to 75 years who were able to read and understand the Sinhala language.

The sample size was computed using the formula for estimating a population proportion of psychological morbidity with absolute precision.<sup>5</sup> The latter was considered to be 30% for Sri Lanka<sup>6</sup> and based on that the computed sample size was 323 for a confidence level of 95% and an absolute precision of 5%.

All patients who were registered for OPD visits in the CNTH (who were eligible and consented to be on the study) during the study period were recruited until the required sample size was reached using convenience sampling. The number recruited per day varied between 10 to 15.

The responses for the 30 question items were rated on a four-point Likert scale (0-1-2-3). A record sheet was used to note down the decision made by the Psychiatrist with regard to presence of depression/anxiety/social dysfunction. This was considered as the reference standard for which the diagnostic criteria were based on ICD 10 classification. Data entry was carried out using EPI INFO 6. Double entry was carried out to check for the reliability of data entry. The data analysis was performed using the statistical package SPSS windows Version 16. The Likert scale was transformed into GHQ score (0,0,1,1) for the analysis of validity. Sensitivity and specificity of the instrument was determined for different cutoff values and the best cutoff value was identified using Receiver Operation Characteristic (ROC) curves which were based on the tradeoff between sensitivity and 1- specificity. The area under the curve (AUC) represents a summary measure of the ability of the GHQ to discriminate between cases and non cases. Sensitivity and specificity were expressed with its 95% confidence intervals (95%CI). Sex and educational and age group differences were also explored during the analysis.

Informed consent was obtained from all the research participants and confidentiality of data obtained were assured by concealing the identity using a coding system. Ethical clearance was obtained from the Ethics Review Committee, Faculty of Medicine, University of Kelaniya. The Psychiatrist took necessary steps to attend to those who required treatment.

## **Results**

Total number of OPD patients who were invited to the study during the study period was 432. Of them 22 patients were not willing to participate which gives a non-participant rate of 5.0 %. Therefore the total patients recruited to the study were 410. Of them 12 (2.9%) were not accessible for psychological assessment by the psychiatrist and 24 (6%) had missing data for one or more items of the GHQ-30. These 36 (9.0%) were excluded from the final analysis, which left a study sample of 374. Of the 374 patients who were assessed by the psychiatrists, 83 (22.2%) had at least one, 27 (7.0%) two and 16 (4.3%) three psychological morbidities. Among the psychiatric diseases identified, 118 (31.6%) suffered from depression, 41 (11%) from anxiety and 26 (7%) from social dysfunction. The percentage of psychological morbidities according to socio-demographic characteristics are presented in Table 1.

Median age of the study population was 32 years (IQR= 21years). The highest proportion (29%; n=112) of them were in the age group of 21 to 30 years. Sixty two percent (n=233) of the study population was Buddhists and 24% (n=90) Catholics or Christians. Two hundred and six (55 %) had studied up to General Certificate Examination (Ordinary Level). The median GHQ score was 3 (IQR=7) for the total sample,, 9 (IQR=12) for those with psychological morbidities and 2 (IQR=5) for normal individuals.

### ***Determination of optimal cutoff level for GHQ 30***

The optimal cutoff level is the point on the curve that has the shortest distance from the left upper corner of the Y axis to the curve. According to the given ROC curve the best cutoff level

was 5/6 for the GHQ-30 with a sensitivity of 67.5% (95% CI: 59% – 76%), a specificity of 80% (95% CI: 75% – 85%) for an AUC of 79% (95% CI: 74% – 84%) [Figure 1]. The GHQ-30 misclassified 26% of subjects either with false positives or false negatives when the cutoff level was 5/6 (Table 2) which gave the lowest misclassification rate. Table 3 shows the best cutoff levels determined by ROC curves for different levels of socio-demographic variables. For all the subgroups considered, the best cutoff level was 5/6 except for females, the older age group and the Catholics/Christians. The AUCs were above 70% for all the subgroups except for Catholics/Christians.

According to the analysis of GHQ score specific (multilevel) likelihood ratios<sup>7</sup>, the latter was observed to have a value of less than one up to a cutoff score of 6 (Table 4) and more than one, at a score of  $\geq 7$ , for which the sensitivity was lower (64% : 95% CI: 56% – 73%) and specificity (84% : 95%CI: 79% – 88.5%) was higher than at 6.

## Discussion

We found that the best cutoff level for the GHQ-30 to detect psychological morbidity was the score of  $\geq 6$ , with a sensitivity of 67.5% and a specificity of 80%. Subgroup analysis of the sample also derived the same cutoff value except for females which was  $\geq 5$ .

Several studies<sup>8,9,10,11</sup> across cultures also reported  $\geq 6$  to be the best cutoff value for GHQ-30. For testing the Greek version<sup>9</sup> of the GHQ-30 included amongst these studies, the setting was an outpatient clinic as well as for ours.

However, the cutoff value suggested by authors of the original GHQ<sup>12</sup> was 4/5 for the GHQ-30. Further they recommended that higher cutoff values than the default, probably would be needed for optimal discrimination between cases and non-cases among physically ill people. We recruited the sample from the OPD of a teaching hospital and majority of them would have had some form of physical illness which explains their presence at the hospital. Therefore, the derived cutoff value of 5/6 for our study is justifiable, as the authors of the GHQ also emphasized for the need to have the cutoff level determined based on research evidence. However, the best cutoff level for the validation of the Italian<sup>10</sup> and the Malay versions<sup>11</sup> of the GHQ-30, which were community based was also 5/6. Further, the above two studies<sup>10,11</sup> demonstrated higher sensitivity than that of our study at this cut off level.

A likelihood ratio of positive test was 3.41 when the cutoff level was set at 5/6. When the cutoff level was set at 6/7, the sensitivity and specificity were 64% and 84% respectively and a misclassification rate of 26% which was the lowest. However, a study from Japan<sup>13</sup> also had recommended a cutoff level of 6/7 for GHQ-30.

We have also analysed stratum specific likelihood ratios which is considered as more useful for bed side diagnosis purposes than the conventional ones for dichotomized outcomes. The latter entails description of cases/ non cases based on a fixed cutoff level, where as the former assesses the likelihood of caseness at each score level given test positive results. This approach also helps

to minimize spectrum bias.<sup>14</sup> The GHQ score specific likelihood ratio was below one (0.88) for scores of  $\geq 6$ , and rose above a value of  $>1$  from a score of 7 onwards. Despite an extensive literature survey, articles reporting GHQ score stratum specific likelihood ratios for GHQ-30 could not be found.

The study participants for the study were recruited without applying any probability sampling method. Even though this might compromise the external validity of the study, the minimal non response rate observed was likely to negate its effect. The selection of cases was not based on prior knowledge of the mental status of the participants. Thus as the sample was not purposively selected, spectrum bias would be minimal for this study and generalizability of the results may be considered high for any primary care setting.

The estimate of prevalence of psychiatric disorders is dependent on the criteria for "casesness" used for validation. This in turn is dependent on the chosen reference standard. Present State Examination (PSE), Clinical Interview Schedule (CIS) and Diagnostic and Statistical Manual of mental disorders, 4<sup>th</sup> edition (DSM-IV) of the American Psychiatric Association and Composite International Diagnostic Interview (CIDI) are few of the reference standards used for the purpose. With regard to the present study the reference standard was based on the psychiatrist's assessment made on uniform standard ICD -10 criteria. The inconsistency observed with regard to the cutoff levels is attributed to the diversity in defining caseness.<sup>15</sup> Goldberg et al<sup>16</sup> reported that CIS is almost identical to that for the PSE. Further they reported that the results based on DMS-IV and the International Classification of Disease 10<sup>th</sup> Edition was comparable.

We agree with Goldberg et al<sup>16</sup> that GHQ can be used efficiently in developing countries as well and that sex, age and educational level do not affect its' validity. Therefore considering the recommendations by the User's Guide on General Health Questionnaire, we give priority to the sensitivity level against specificity to set the threshold level at 5/6 of GHQ-30 for case detection.<sup>13</sup> Otherwise the threshold level of 6/7 is recommended for GHQ-30. In conclusion GHQ 30 displays adequate validity for use in the Sinhala speaking primary care setting attendees for assessing psychiatric disorders.

### **Acknowledgements**

This study was funded by the University of Kelaniya. We are grateful to the Research and Publication Committee to the University of Kelaniya and to the research assistants.

**Conflict of Interest:** None declared.

---

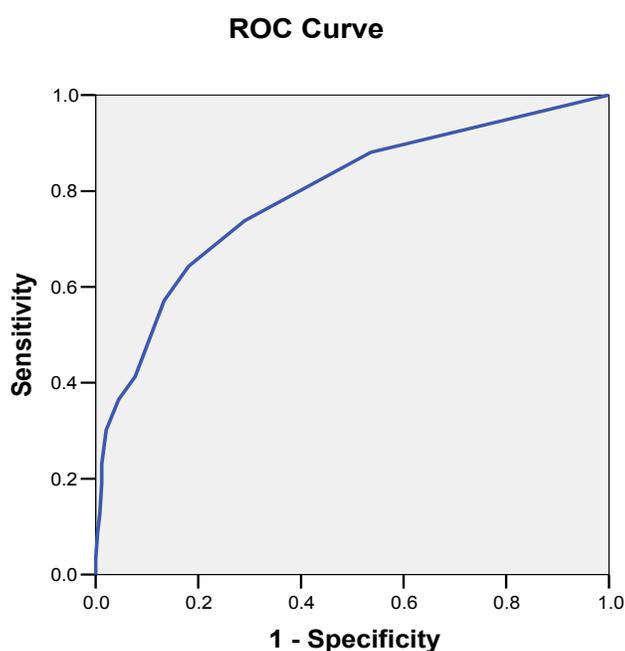
### **References**

1. Sriram TG, Chandrashekar CR, Isaac MK, Shanmugham V. The General Health Questionnaire (GHQ) Comparison of the English version and a translated Indian version. *Social Psychiatry and Psychiatric Epidemiology*. 1989;24:317-320.

2. Abramson JH, Abramson ZH. Survey methods in community medicine. London, Churchill Livingstone;1990.
3. Huppert FA, Walters DE, Day NE, Elliott BJ. The factor structure of the General Health Questionnaire (GHQ-30). A reliability study on 6317 community residents. *The British Journal of Psychiatry* 1989;155: 178-185.
4. Jakob KS, Bhugra D, Mann H. The validation of the 12-item General Health Questionnaire among ethnic Indian women living in the United Kingdom. *Psychological Medicine*. 1997;27:1215-1217
5. Lwanga SK, Lemeshow S. Sample size determination in health studies. A practical manual. WHO 1991.
6. National Mental Health Survey, 2007. Ministry of Health, Sri Lanka.
7. Fletcher RH, Fletcher SW, Wagner EH. Clinical epidemiology: the essentials. 3<sup>rd</sup> edition Williams and Wilkins, USA; 1996.
8. Montazeri A, Harirchi AM, Shariati M, Garmaroudi G, Ebadi M, Fateh A. The 12-item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version. *Health and Quality of Life Outcomes* 2003, 1:66
9. Garyfallos G, Karastergiou A, Adamopoulou A, Moutzoukis C, Alagiozidou E, Mala D, Garyfallos A. Greek version of the General Health Questionnaire: accuracy of translation and validity. *Acta Psychiatr Scand*. 1991 Oct;84(4):371-8.
10. Febbo S, Burvill PW. Validation of an Italian translation of the 30-item General Health Questionnaire for use in Australia. *Aust N Z J Psychiatry*. 1995 Jun;29(2):266-9.
11. Yusoff MSB. The Validity of Two Malay Versions Of The General Health Questionnaire (GHQ) In Detecting Distressed Medical Students. *ASEAN Journal of Psychiatry*.2010;11(2): <http://www.aseanjournalofpsychiatry.org/oe11201.htm>.
12. Goldberg D, Williams P. A user's guide to the General Health Questionnaire, Windsor:NFER-Nelson,1991
13. Itawa N. A brief introduction to the self-administered questionnaires for minor psychiatric impairment. *Arch psyciat Diag Clin Eval*. 1992;3:413-427
14. N. Schmitz, J. Kruse, W. Tress. Application of stratum-specific likelihood ratios in mental health screening *Social Psychiatry and Psychiatric Epidemiology*. Vol 35 (8): 375-379. DOI: 10.1007/s001270050253"

15. Goldberg DP, Oldhinkel T and Ormel J: Why GHQ threshold varies from one place to another. *Psychol Med* 1998, 28:915-921.

16. Goldburg DP et al. Validity of two versions of the GHQ in the WHO study of Mental illness in general health care. *Psycholl Med*.1997; 27:191 – 197.



**Fig. 1** The fitted ROC curve for the GHQ-30

**Table 1: Psychological morbidities by socio-demographic characteristics**

	Psychological morbidity (Reference standard)				
	Yes		No		
	n	%	n	%	
<b>Gender</b>					
Male (n=122)					
Female (n=225)	44	(36%)	78	(64%)	
	75	(33.3)	150	(66.7%)	
<b>Age (years)</b>					
<32 (n=191)					
≥ 32 (n=182)	55	(28.8)	136	(71.2)	
	70	(38.5)	112	(61.5)	

<i>Educational Level</i>					
<grade 10 (n=206)					
≥grade 10 (n=155)	80	(38.8)	126	(61.2)	
	41	(26.5)	114	(73.5)	
<i>Income</i>					
Low (n=99)					
High (n=101)	40	(40.4)	59	(59.6)	
	26	(25.7)	75	(74.3)	
<i>Religion</i>					
Buddhist (n=233)					
Catholics (n=90)	75	(32.2)	158	(67.8)	
	31	(34.4)	59	(65.6)	

**Table 2: Sensitivity, specificity, misclassification rate, likelihood ratio and psychological morbidity for selected cutoff values of GHQ 30**

GHQ score	Reference Standard		Sensitivity	Specificity	Misclassification rate	Likelihood Ratio
	Yes n (%)	No n (%)				
≥1	116 (92.1)	167 (67.3)	<b>92.1</b>	<b>32.7</b>	<b>37.6</b>	<b>1.37</b>
≥2	109 (86.5)	132 (53.2)	<b>86.5</b>	<b>46.8</b>	<b>33.3</b>	<b>1.62</b>
≥3	104 (82.5)	111 (44.8)	<b>82.5</b>	<b>55.2</b>	<b>31.1</b>	<b>1.84</b>
≥4	97 (77.0)	85 (34.3)	<b>77.0</b>	<b>65.7</b>	<b>28.6</b>	<b>2.24</b>
≥5	90 (71.4)	64 (25.8)	<b>71.4</b>	<b>74.2</b>	<b>27.2</b>	<b>2.77</b>
≥6	85 (67.5)	49 (19.8)	<b>67.5</b>	<b>80.2</b>	<b>26.1</b>	<b>3.41</b>
≥7	81 (64.3)	40 (16.1)	<b>64.3</b>	<b>83.9</b>	<b>25.9</b>	<b>3.99</b>
≥8	74 (58.7)	33 (13.3)	<b>58.7</b>	<b>86.7</b>	<b>27.3</b>	<b>4.41</b>
≥9	66 (52.4)	26 (10.5)	<b>52.4</b>	<b>89.5</b>	<b>29.0</b>	<b>4.99</b>
≥10	60 (47.6)	24 (9.7)	<b>47.6</b>	<b>90.3</b>	<b>31.0</b>	<b>4.90</b>
≥11	54 (42.9)	19 (7.7)	<b>42.9</b>	<b>92.3</b>	<b>32.4</b>	<b>5.57</b>
≥12	46 (36.5)	14 (5.6)	<b>36.5</b>	<b>94.4</b>	<b>30.5</b>	<b>6.52</b>
≥13	44 (34.9)	12 (4.8)	<b>34.9</b>	<b>95.2</b>	<b>35.0</b>	<b>7.27</b>
≥14	41 (32.5)	06 (2.4)	<b>32.5</b>	<b>97.6</b>	<b>35.0</b>	<b>13.5</b>
≥15	35 (27.8)	04 (1.6)	<b>27.8</b>	<b>98.4</b>	<b>37.0</b>	<b>17.3</b>
<b>Total</b>	126 (100)	248 (100)				

**Table 3: Optimal cutoff values, sensitivity, specificity and area under the curve (AUC) of the Sinhala version GHQ 30 by different levels of socio-demographic variables**

	Best cutoff value	Sensitivity	Specificity	Area under the curve	95% CI for AUC
<i>Gender</i>					
Male (n=122)	5/6	0.77	0.74	0.81	0.72 – 0.89
Female (n=225)	4/5	0.67	0.79	0.79	0.73 – 0.86
<i>Age (years)</i>					
<32 (n=191)	5/6	0.71	0.76	0.79	0.71 – 0.87
≥ 32 (n=182)	4/5	0.70	0.80	0.80	0.73 – 0.87
<i>Educational Level</i>					
<grade 10 (n=206)	5/6	0.67	0.80	0.80	0.74 – 0.86
≥grade 10 (n=155)	5/6	0.68	0.80	0.78	0.69 – 0.87
<i>Income</i>					
<3750 SLR (n=99)	5/6	0.65	0.78	0.77	0.68 – 0.87
≥3750 SLR (n=101)	5/6	0.81	0.80	0.88	0.81 – 0.96
<i>Religion</i>					
Buddhist (n=233)	5/6	0.70	0.83	0.83	0.77 – 0.88
Catholics (n=90)	4/5	0.74	0.68	0.73	0.62 – 0.84

SLR; Sri Lankan rupees, CI; confidence interval, AUC; area under the curve

**Table 4: Multilevel likelihood ratios for GHQ 30**

GHQ score	Reference Standard				Likelihood Ratio
	n	Yes (%)	No n	(%)	
0	10	(07.9)	81	(32.7)	<b>0.24</b>
1	07	(05.6)	35	(14.1)	<b>0.40</b>
2	05	(04.0)	21	(08.5)	<b>0.47</b>
3	07	(05.6)	26	(10.5)	<b>0.53</b>
4	07	(05.6)	21	(08.5)	<b>0.66</b>
5	05	(04.0)	15	(6.0)	<b>0.66</b>
6	04	(03.2)	09	(03.6)	<b>0.88</b>
7	07	(05.6)	07	(02.8)	<b>2.00</b>
8	08	(06.3)	07	(02.8)	<b>2.25</b>
9 – 15	34	(26.9)	22	(08.87)	<b>3.00</b>
16 – 30	30	(23.8)	04	(01.61)	<b>14.8</b>
<b>Total</b>	126	(100)	248	(100)	