THE VALIDATION OF THE GENERAL HEALTH QUESTIONNAIRE (GHQ-28) IN A PRIMARY CARE SETTING IN SAUDI ARABIA

Abdulrazzak Alhamad, *MD,[*]* ***Eiad A. Al-Faris,****MRCGP[†]*
Departments of [*]Psychiatry and [†]Family and Community Medicine, College of Medicine, King Saud University, Riyadh, Saudi Arabia

---

هدف الدراسة: الهدف من هذه الدراسة هو تقنين الترجمة العربية لاستبانة الصحة العامة-28 في عيادات الرعاية الصحية الأولية بالمملكة العربية السعودية.
طريقة الدراسة: تم اختيار ستين مريضا سعوديا بطريقة عشوائية منظمة حيث طلب منهم تعبئة نموذج استبانة الصحة العامة – 28 ثم تم فحصهم من قبل استشاري الطب النفسي باستخدام الترجمة العربية لبيان المقابلة السريرية.
نتائج الدراسة: لقد وجد أن الحد الأفضل للاستبانة عند أعلى قيمتين لحساسية الاستبانة ودقته هي 5/4 ووجد أن معايير التقنين عند هذا الحد هي الحساسية (72%)، الدقة (74%) مقدار التشخيص الإيجابي (72%) ، مقدار التشخيص السلبي (74%) ونسبة خطأ التشخيص (27%). وقد وجد معامل الارتباط (+ 0.61) ومعامل ارتباط سبيرمان (+ 0.57) ووجد أن المساحة تحت منحنى روك (69%). إن نتائج هذه الدراسة في الحد الأفضل مشابهة لغيرها من الدراسات في الرعاية الصحية الأولية. وبالرغم من أن معايير التقنين منخفضة نسبيا، إلا أنها في حدود الذي وجدته الدراسات الأخرى سواء في الولايات المتحدة أو بريطانيا أو الدول الأخرى النامية. وهذا يؤكد الاقتراح بضرورة إيجاد استبيان للصحة العامة خاص بالبيئة العربية يقوم على النسخة المترجمة لاستبيان الصحة العامة مع إضافة أسئلة تختص بالبيئة للمساعدة في تشخيص المرض النفسي.
الخلاصة: إن النسخة العربية لاستبيان الصحة العامة-28 تعتبر مقننة على البيئة السعودية في عيادات الرعاية الصحية الأولية ويمكن أن تساعد أطباء الرعاية الأولية كثيرا على زيادة اكتشاف الإضرابات النفسية وكذلك في الأبحاث المسحية لها.
الكلمات المرجعية: استبيان، الصحة العامة، الرعاية الأولية، المملكة العربية السعودية.

Objective: *The objective of this study was to validate an Arabic version of the General Health Questionnaire (GHQ-28) in a primary care setting in Saudi Arabia.*
Methodology: *A total of 60 Saudi patients selected by means of systematic random sampling were asked to fill out the GHQ-28 Arabic version. The psychiatrist interviewed all patients using the Arabic version of the Clinical Interview Schedule (CIS).*
Results: *The best cut-off level for the GHQ-28 indicating best trade-off between sensitivity and specificity was 4/5, where the validity values were, sensitivity; 72%, specificity; 74%, positive predictive value; 72%, negative predictive value; 74% and misclassification rate; 27%. The correlation coefficient was r = +0.61 and the Spearman's Rank-difference correlation was rs = +0.57. The area under the ROC Curve was 69%. The cut-off point 4/5 in this study is the same as recommended by others in primary care settings. Although the validity parameters are relatively low,*

Correspondence to:
Dr. Abdulrazzak Alhamad, *Assistant Professor and Consultant Psychiatrist, Head, Division of Psychiatry, College of Medicine, King Saud University, P.O. Box 7805, Riyadh 11472, Saudi Arabia*

they are within the range found by other studies in USA, UK and developing countries. This supports the suggestion to develop an Arabic Screening Questionnaire based on the translated GHQ with the addition of culturally specific items.
Conclusion: *The GHQ-28 Arabic Version is a valid instrument that may be of great help to primary care doctors in improving detection of psychiatric morbidity and in epidemiological research.*

Key Words: *General Health Questionnaire, Primary Care, Saudi Arabia*

INTRODUCTION

The General Health Questionnaire (GHQ) in its full 60-items, or abbreviated 30 and 28 items version is by far the most popular screening instrument.[1] It is a self-reporting questionnaire developed by Goldberg (1970), to detect functional psychiatric disorders in the community and primary care settings.[2,3] It has been extensively tested in various cultures and linguistic groups in primary care and other settings generally showing good validity results.[4-9] Tarnopolsky et al suggested that the GHQ should be standardized on the population where it is to be applied, because validity coefficients obtained in one setting do not necessarily hold in another.[4] Psychiatric disorders are shown to form a major part of morbidity in Saudi primary health care, and the vast majority of cases are missed.[10,11] Therefore, validation of a screening instrument such as the GHQ in the patients' language takes a little time, and is important in epidemiological research to improve detection and recognition of psychiatric morbidity.[3,4] A medline and a Saudi literature search showed no study that validated the GHQ in primary health care in Saudi Arabia.

The objective of the present study was to validate an Arabic version of the GHQ-28 (see appendix) in primary care setting in Saudi Arabia against the Psychiatrist's assessment by means of the Clinical Interview Schedule (CIS).

METHODOLOGY

Subjects: The study was carried out in the primary care center attached to King Abdulaziz University Hospital. It is situated in the center of Riyadh and serves mostly a Saudi population of different social classes. The study population included patients of both sexes above 14 years of age attending the primary care clinics for any reason.

A total of 60 Saudi patients selected by means of systematic random sampling were asked to fill out the GHQ-28 Arabic version while waiting to be seen by their doctor. Patients who were found to be illiterate or experienced difficulty in filling out the questionnaire were helped by a trained nurse.

All patients were interviewed by the first author, a consultant psychiatrist with eleven years post-qualification experience, using the Arabic version of the CIS. Each patient was given a score on the psychiatric severity rating (0-4) as follows: 0=no psychiatric disturbance, 1=mild subclinical psychiatric disturbance, 2=clinically significant (mild) psychiatric disturbance, 3=clinically significant (moderate) psychiatric disturbance, 4=clinically significant (marked) psychiatric disturbance.

Instruments: For the present study the instruments CIS and GHQ-28 were translated into Arabic by two Arab psychiatrists. The reliability of the Arabic versions was checked with a translation into English by another psychiatrist who had no knowledge of the instrument and they were found to be in close agreement. The underlying assumption

is that the psychiatrist's assessment using the CIS is the gold standard against which the GHQ-28 is compared. The CIS is divided into four sections containing ratings based on symptoms reported by the patient (arranged in ten groups), twelve items that represent the interviewer's view of the manifest abnormalities and an ICD clinical diagnosis. The scoring of the GHQ-28 depends on the response category the patient chooses for each of the 28-items and the scoring developed by Goldberg (1970) was used to count responses in codes 3 and 4 only. The GHQ-28 version was chosen for the study because it was short and was found to be more valid than both the GHQ-12 and the GHQ-30.[12]

Validation parameters: The validity of the GHQ-28 was estimated by the following: evaluating the sensitivity and specificity at best trade-off point; estimating the positive predictive value, the negative predictive value and the misclassification rate; determining the simple correlation coefficient (r) and the Spearman's Rank-Difference Correlation values; finding the confidence intervals between observed data values of GHQ-28 and its predicted data values in a scattergram; and using the Receiver (relative) operating characteristic (ROC) analysis. The ROC curve is a graphic representation of the relationship between sensitivity and specificity for a diagnostic instrument. It is constructed by plotting the sensitivity (true positive rate) against the false positive rate (1-specificity) for all possible GHQ-28 cut-off points. The resulting plot, which takes the shape of a curve, is known as a ROC curve (figure 1).[13]

RESULTS

A total of sixty patients, 30 males and 30 females were included in the study. The age range was 14-70 and the mean age was 28.1± 10.77. The mean age for males was 29.17 ± 10.99 and for females was 27.1 ± 10.63.

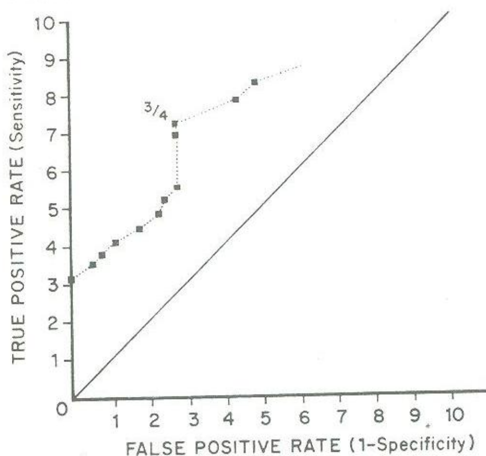

Figure 1:ROC curve of the GHQ-28

Table 1 shows different cut-off levels of the GHQ-28 (2/3) through (20/21) and their comparable sensitivity, specificity, positive predictive values and negative predictive values. Patients who scored above these threshold levels were considered as probable cases while those who scored below them were considered probable normals. The overall best cut-off level indicating the best diagnostic ability of the GHQ-28 was found to be 4/5, where the misclassification rate was 27%, the sensitivity was 72%, the specificity was 74%, the percentage of cases missed was 14%, the positive predictive value was 72% and the negative predictive value was 74%. The validity values for males and females respectively were sensitivity 64% and 80%, specificity 81% and 66.9% and misclassification rate was 26.7% for both sexes.

The ROC curve of the GHQ-28 (Figure 1), showed that it was a valid test and the area under the curve measured graphically was 69%. When the GHQ-28 scores were plotted against the CIS overall severity rating scores, the correlation coefficient was r=+0.61 and the p-value approached zero; the Spearman's Rank-difference correlation was rs=+0.57 and

Table 1: GHQ-28 threshold scores and their validity parameters

Threshold score	Sensitivity %	Specificity %	+ve predictive value %	-ve predictive value %
2/3	83	52	62	76
3/4	79	58	64	75
4/5	72	74	72	74
5/6	71	74	71	74
6/7	70	75	71	74
7/8	55	76	55	64
8/9	52	77	69	63
9/10	48	81	70	63
10/11	45	84	72	62
11/12	41	90	80	62
12/13	38	94	85	62
13/14	35	97	91	61
14/15	31	100	100	61
15/16	21	100	100	57
16/17	21	100	100	57
17/18	14	100	100	55
18/19	14	100	100	55
19/20	14	100	100	55
20/21	10	100	100	54

the p-value approached zero which is highly significant (figure 2). The scattergram of data and regression line were also found to be highly significant and the confidence limit was 99.0% (figure 3).

DISCUSSION

This study has shown that the GHQ-28 is a valid and useful screening instrument of psychiatric morbidity in a primary care setting in Saudi patients. The best cut-off point was found to be 4/5 which is comparable to what was recommended by Goldberg for general practice settings.[3] This cut-off point represented the best trade-off between sensitivity (72%) and specificity (74%). Although, the validity of the GHQ-28 in our study was not high, it was still in the range reported in many other studies.[4,9,14,15] Even though, Goldberg[9] had higher sensitivity and specificity values, other studies conducted in British and United States communities had lower values for sensitivity and specificity comparable to our rsults.[15-17] Table 2 shows data comparing different validation studies of the

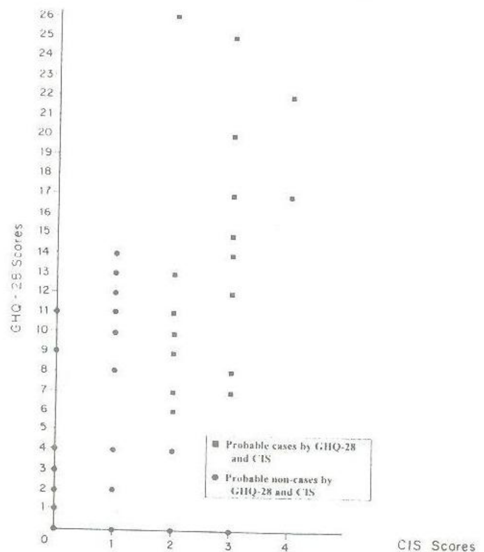

Figure 2:GHQ-28 and CIS scores

Table 2: GHQ-28 different validation studies

Study	Cut-off point	Sensitivity %	Specificity %	Correlation coefficient (r)	Misclassification rate
Goldberg (1972)	4/5	91	87	+0.77	11
Goldberg & Hillier (1979)	4/5	88	84	+0.76	14.5
Tarnopolsky et al (1979)	4/5	78	72	+0.45	26
Present study	4/5	72	74	+0.61	27

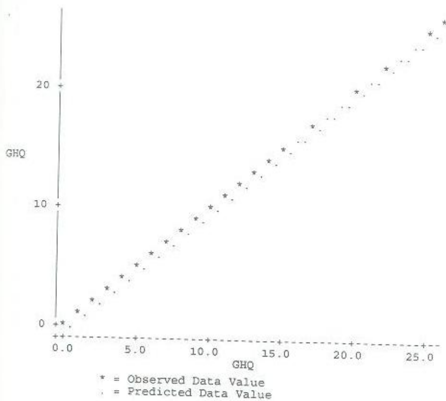

***Figure 3:** Scattergram of data & regression line & confidence intervals*

GHQ-28 in the primary care setting. The relatively lower value of sensitivity for males and specificity for females is consistent with other studies.[18,19] On the other hand, since previous studies have not shown strong associations between sex and GHQ validity,[2] our results might show that a higher cut-off level may be used for females. This may be explained by the tendency of Saudi females to somatise.[13,20] Women in general tend to report more symptoms[5] and score high on the GHQ, thus attaining false-positive rates, and reducing the specificity.

The validity of the GHQ-28 was also shown to be very good by the ROC curve, which had been used previously in the validation of the GHQ-28.[5,13,18]

The use of longer GHQ versions may increase validity,[2,3] but shorter ones such as the GHQ-28 are still valid and take much less time. Primary care physicians miss about 50% of psychiatric cases[10,16,21] and the use of the GHQ-28 may be useful to improve detection rate. Goldberg suggests that when a patient is found to have a high score, the most natural response by the clinician is to look at the questionnaire again with the patient and ask additional probing questions suggested by particular symptoms.[3,16] It is also important to note that GHQ-28 has been found to have an important role in alerting primary care doctors not to miss new episodes of psychiatric morbidity in patients with chronic physical diseases.[16]

Hence, GHQ-28 is useful in epidemiology.[3,4,16,18] The present validation study is also important for researchers who intend to estimate the prevalence of psychiatric disorders in primary care in the Saudi communities and plan services to improve recognition and diagnosis of psychiatric morbidity in primary care.

The fairly low validity in our study conforms with other validity studies in developing countries and may strongly support the suggestion by other researches[22,23] to develop and validate a new Arabic screening questionnaire that builds on the translated GHQ

with the addition of specific items relevant to the Saudi culture for the identification of psychiatric morbidity.

REFERENCES

1. Mayou R, Hawton K. Psychiatric disorders in the General Hospital. British J Psych 1986;149:172-90.
2. Goldberg D. Manual of the General Health Questionnaire. Windsor (UK): National Foundation for Educational Research; 1978.
3. Goldberg D. Use of the General Health Questionnaire in Clinical Work. BMJ 1986;293:1188-9.
4. Tarnopolsky A, Hand DJ, Mclean EK. Validity and uses of a screening questionnaire (GHQ) in the community. British J Psych 1979;134:508-15.
5. Bridges KW, Goldberg DP. The validation of the GHQ-28 and the use of the MMSE in neurological inpatients. British J Psych 1986;148:548-53.
6. Mari JJ, Williams P. Minor psychiatric disorder in primary care in Brazil: A pilot study. Psychological Medicine 1984;14:223-72.
7. Cheng TA. A pilot study of mental disorders in Taiwan. Psychological Medicine 1985;15:195-204.
8. Sato T, Takeichi M. Lifetime prevalence of specific disorders in a general medicine clinic. General Hospital Psychiatry 1993;15:224-33.
9. Goldberg DP, Blackwell B. Psychiatric illness in general practice: a detailed study using a new method of case identification. BMJ 1970;2:439-43.
10. Al-Faris EA, Alhamad A, Al-Shammari S. Hidden and conspicuous psychiatric morbidity in Saudi primary health care (A pilot study). The Arab J of Psychiatry 1995;6:162-75.
11. Al-Faris E, Al-Subaie A, Khoja T, Al-Ansary L, Abdul-Raheem F, Al-Hamdan N, et al. Training primary health care physicians in Saudi Arabia to recognise psychiatric illness. Acta Psychiatr Scand 1997;96:439-44.
12. Banks MH. Validation of the General Health Questionnaire in a young community sample. Psychological Medicine 1983;13:349-53.
13. Ly Kouras, Adrachta D, Kalfakis N, Oulis P, Voulgani A, Christodoulu GN, et al. GHQ-28 as an aid to detect mental disorders in neurological inpatients. Acta Psychiatr Scand 1996;93:212-6.
14. Goldberg DP. The detection of psychiatric illness by questionnaire. London: Oxford University Press; 1972.
15. Goldberg D. Identifying psychiatric illness among general medical patients. BMJ 1985; 291:161-2.
16. Wright AF, Perini AF. Hidden psychiatric illness: use of the general health questionnaire in General Practice. Journal of the Royal College of General Practitioners 1987;37:164-7.
17. Cleary PD, Goldberg ID, Kessler IG, Nyez GR. Screening for mental disorder among primary care patients. Archives of general psychiatry 1982;39:837-40.
18. Stanfeld SA, Marmot MG. Social class and minor psychiatric disorder in British civil servants: a validated screening survey using the GHQ. Psychological Medicine 1992;22:739-49.
19. Hobbs P, Ballinger CB, Greenwood C, Martin B, McClure A. Factor analysis and validation of the General Health Questionnaire in men: A general practice survey. British J Psych 1984;144:270-5.
20. Racy J. Somatisation in Saudi women: a therapeutic challenge. British J Psych 1980; 137:212-6.
21. Goldberg D, Steele JJ, Johnson A, Smith C. Ability of Primary Care Physicians to make accurate ratings of psychiatric symptoms. Archives of General Psychiatry 1982;39:829-33.
22. Fontanesi F, Gobetti C, Zumermann-Tansella CH, Tansella M. Validation of the Italian version of the GHQ in a general practice setting. Psychological Medicine 1985;15:411-15.
23. Cheng T, Williams P. The design and development of a screening questionnaire (CHQ) for use in community studies of mental disorders in Taiwan. Psychological Medicine 1986;16:415-22.

Appendix

استبيان عن الصحة العامة

الرجاء قراءة مايلي بعناية :-

نود أن نعرف إن كان لديك أى شكوى مرضية (علة) وكيف كانت حالتك الصحية العامة خلال الاسابيع القليلة الماضية .

نرجو الاجابة على جميع الاسئلة التالية وذلك بوضع خط تحت الاجابة المناسبة التي هى أقرب انطباق (شبيهة) على حالتك الصحية.

نرجو أن تتذكر أننا نود معرفة الشكاوى المرضية والصحية التى تعانى منها الآن والتى عانيت منها خلال الاسابيع القليلة الماضية فقط وليست التى عانيت منها في الماضى البعيد .

من المهم الاجابة على كل الاسئلة ، شكراً على حسن تعاونك .

السؤال	١	٢	٣	٤
١- هل كنت تتمتع بصحة تامة وجيدة ؟	احسن من المعتاد	لافرق كالمعتاد	أسوأ من المعتاد	أسوأ بكثير من المعتاد
٢- هل تشعر بحاجة إلى ادوية مقوية .	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٣- هل تشعر انك منهك (تعبان) وفي حالة غير طبيعية؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٤- هل شعرت (حسيت) انك مريض ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٥- هل شعرت بأى آلام في رأسك ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٦- هل شعرت بأى شد في رأسك ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٧- هل شعرت بنوبات سخونة (حرارة) أو برودة ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٨- هل قل نومك بسبب القلق ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٩- هل تجد صعوبة في مواصلة النوم ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
١٠- هل شعرت انك تحت ضغط باستمرار ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
١١- هل اصبحت حاد الطبع وسريع الأنفعال ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
١٢- هل ينتابك خوف أو رعب بدون سبب مقبول ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
١٣- هل تحس أن كل ما حولك اصبح عبئا عليك ؟	لا أبـــداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
١٤- هل تشعر انك متوتر الاعصاب ومتحفز كل الاوقات؟	لا بالتأكيد	ليس اكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
١٥- هل كان في استطاعتك الاستفادة من وقتك بالصورة المطلوبة ؟	اكثر من المعتاد	كالمعتاد	اقل من المعتاد	اقل بكثير من المعتاد
١٦- هل اصبحت تحتاج لوقت اطول لانجاز اغراضك العادية ؟	اسرع من المعتاد	كالمعتاد	اطول من المعتاد	اطول بكثير من المعتاد
١٧- هل حسيت انك كنت تؤدى اعمالك بصورة جيدة	احسن من المعتاد	كالمعتاد	اقل من المعتاد	اقل بكثير من المعتاد
١٨- هل انت راضٍ عن الطريقة التى انجزت بها مهامك؟	اكثر رضا	راضى كالمعتاد	اقل رضا من المعتاد	اقل رضا من المعتاد
١٩- هل شعرت انك تقوم بدور مفيد لما حولك ؟	اكثر من المعتاد	كالمعتاد	اقل فائدة من المعتاد	اقل فائدة بكثير من المعتاد
٢٠- هل كان في استطاعتك اتخاذ القرار في الامور ؟	اكثر من المعتاد	كالمعتاد	اقل من المعتاد	اقل استطاعة بكثير
٢١- هل كنت تجد متعة في اداء نشاطك ؟	اكثر من المعتاد	كالمعتاد	اقل من المعتاد	اقل بكثير من المعتاد
٢٢- هل كنت تنظر لنفسك كشخص عديم الفائدة ؟	لا مطلقاً	ليس اكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٢٣- هل شعرت ان حياتك ميئوس منها ؟	لا أبـداً	ليس اكثر من المعتاد	اكثر من المعتاد	اكثر بكير من المعتاد
٢٤- هل شعرت أن الحياة لاقيمة لها ؟	لا أبـداً	ليس أكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٢٥- هل فكرت ان تنهى حياتك ؟	لا بالتأكيد	لا اعتقـد	خطر لى ذلك	نعم بالتأكيد
٢٦- هل وجدت نفسك في بعض الاوقات لاتستطيع عمل شيء لان اعصابك متوترة ؟	لا مطلقاً	ليس اكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٢٧- هل تمنيت لو كنت ميتا وبعيدا عن كل ذلك ؟	لا مطلقاً	ليس اكثر من المعتاد	اكثر من المعتاد	اكثر بكثير من المعتاد
٢٨- هل تفكر في الانتحار باستمرار ؟	لا بالتأكيد	لا اعتقد ذلك	خطرت لى الفكرة	نعم بالتأكيد

A [ ] B [ ] C [ ] D [ ] Total [ ]