Screening for affective and anxiety disorders in medical patients: comparison of HADS, GHQ-12 and Brief-PHQ

Screening nach affektiven und Angststörungen bei Patienten in der medizinischen Versorgung: ein Vergleich von HADS, GHQ-12 und Brief-PHQ

Abstract

Objectives: The detection of patients with comorbid mental disorders is of high clinical importance. Screening instruments can be recommended for early recognition. This study investigates the discriminant validity of the Hospital Anxiety and Depression Scale (HADS), General Health Questionnaire (GHQ-12) and Patient Health Questionnaire (Brief-PHQ-D) in chronically ill patients.

Methods: Two hundred and four patients with chronic illness participated in a two-stage survey. First patients were screened with HADS, Brief-PHQ and GHQ-12 and then examined for mental disorders by clinical standardized interview (M-CIDI). Validity of the three instruments was compared using ROC(receiver operating characteristics)-analysis and CIDI-diagnoses as criteria.

Results: The Brief-PHQ and the HADS performed better than GHQ-12 concerning affective and anxiety disorders without reaching significance. Even though the Brief-PHQ performs significantly better in the category of "any mental disorder", the differences between the Brief-PHQ and the HADS remain not significant considering anxiety and affective disorders. The Brief-PHQ performed slightly better considering depressive disorders with an area under the curve (AUC) of 0.844, a sensitivity of 80% and a specifity of 75.7%.

Conclusions: Screener can be used for the detection of mental disorders in patients with chronical illness. Especially the Brief-PHQ and the HADS can be recommended considering sensitivity and specifity. An advantage of the Brief-PHQ is the ability of categorial and dimensional analysis.

Keywords: screening, affective disorders, anxiety disorders, mental disorders, somatic disorders, epidemiology, comorbidity

Zusammenfassung

Zielsetzung: Das frühzeitige Erkennen komorbider psychischer Störungen bei Patienten mit chronischen Erkrankungen ist von hoher klinischer Bedeutung. Screening-Instrumente können hierbei eine wichtige Hilfestellung sein. Allerdings sollte bei der Auswahl der Instrumente besonderes Augenmerk auf die Validität für den jeweiligen Indikationsbereich gelegt werden. Diese Studie vergleicht die Erkennensleistung der Hospital Anxiety and Depression Scale (HADS), des General Health Questionnaire (GHQ-12) und des Patient Health Questionnaire (Brief-PHQ-D) an einer Stichprobe von Patienten mit chronischen körperlichen Erkrankungen.

Methodik: 204 Patienten aus der medizinischen Rehabilitation nahmen an einer zweistufigen Untersuchung teil : (1) Screening mit HADS, GHQ-12, Brief-PHQ und (2) diagnostisches Interview (M-CIDI). Die Validität der Screeninginstrumente wurde anhand der ICD-10 Diagnosen aus dem CIDI mittels ROC(receiver operating characteristics)-Analysen verglichen.

Daniela Hahn[1]
Katrin Reuter[2]
Martin Härter[2]

1 University of Northampton, School for Social Sciences, Department for Psychology, Northampton, UK

2 Dept. of Psychiatry and Psychotherapy, Section of Clinical Epidemiology and Health Services Research, Freiburg University Medical Centre, Freiburg, Germany

Ergebnisse: Es zeigte sich eine Überlegenheit des Brief-PHQs und der HADS gegenüber dem GHQ im Erkennen von affektiven und Angststörungen. Obwohl der Brief-PHQ signifikant besser in der Kategorie „irgendeine psychische Störung" abschneidet, bleiben die Unterschiede zwischen dem Brief-PHQ und der HADS die anderen Bereiche betreffend nicht signifikant.

Der Brief-PHQ zeigt eine etwas bessere Erkennensleistung hinsichtlich affektiver Störungen mit einem AUC(area under the curve)-Wert von 0,844, einer Sensitivität von 80% und einer Spezifität von 75,7%.

Fazit: Screener sind zum frühzeitigem Erkennen psychischer Störungen bei Patienten mit chronisch körperlichen Erkrankungen gut geeignet. Insbesondere der Brief-PHQ und die HADS können aufgrund ihrer guten Sensitivitäts- und Spezifitätswerten empfohlen werden. Ein Vorteil des Brief-PHQ liegt in seiner Möglichkeit für die kategoriale und dimensionale Auswertung.

Schlüsselwörter: Screening, chronische Erkrankung, psychische Belastung, psychische Störung

Introduction

Over the past years mental disorders have received increased attention in patients with chronic somatic diseases. Studies have shown that patients suffering from chronic diseases have higher prevalence rates of mental disorders than the general population without somatic illness, with affective and anxiety disorders being most common [1], [2], [3], [4], [5].

It is of high clinical importance to recognize psychological symptoms, ranging from psychological burden or strain to mental disorders or psychiatric illnesses that need treatment, early and to introduce psychological or psychiatric intervention in time as to prevent (1) chronification of disorders, (2) negative impact on coping skills and decrease of quality of life as well as (3) raised health care costs due to increased utilization of medical services [6], [7], [8], [9], [10], [11], [12], [13], [14]. Results considering recognition of mental disorders in chronically ill rehabilitation patients though suggest, that there are problems in the detection of mental disorders [13], [15]. Härter et al. found detection rates for rehabilitation patients varying between 32%-48% [13], [15], whereas detection rates in primary care are somewhat higher, ranging between 54%-74% [16], [17], [18], [19], [20], [21], [22]. Problems in detecting mental disorders in somatically ill patients can be explained by overlap of somatic and psychiatric illnesses, a tendency to underestimate negative influence of psychological symptoms in patients with somatic diseases as well as a lack of knowledge about psychological symptoms and syndromes.

Screening instruments can improve recognition rates of mental strain and disorders while being time and cost efficiently. These however must be tested considering their validity for special groups of patients. Also it is important to trade off sensitivity and specificity due to the needs of the user. One of the most used screening instruments is the Hospital Anxiety and Depression Scale

(HADS) [23]. This screener has been especially developed for and tested in patients with somatic diseases. Another well known and widespread screener is the General Health Questionnaire in its twelve item version (GHQ-12) [24]. It is designed as a self report instrument for the measurement of global burden due to disease. Up to now it has been used in many epidemiological studies concerning mental disorders [25], [26], [27], [28], [29]. A relatively new instrument is the Patient Health Questionnaire (Brief-PHQ) [30], which has mainly been used in clinical studies, even though there are not many international studies using it up to now [31].

The discriminating validity of the HADS and GHQ-12 has been tested by our group in patients with cancer, orthopaedic and cardiovascular diseases [32], [33], [34], [35]. In these studies the HADS performed better in recognizing mental disorders than the GHQ-12 did. Nevertheless, using the single scales of the HADS for the recognition of anxiety and affective disorders did not bring any advantage compared to the overall performance [34].

The discriminating validity of Brief-PHQ and HADS in the detection of depressive and anxiety disorders has been tested so far in several clinical studies [36], [37], [38]. In all these studies the Brief-PHQ was significantly superior to the HADS considering the detection of mental disorders.

Up to now a comparison of discriminating validity between the three screeners HADS, GHQ-12 and Brief-PHQ for patients with somatic diseases is missing. Because the three instruments differ concerning their methodological conception (dimensional vs. categorical and dimensional approach), were used in different settings until now (clinical and epidemiological settings) and have different specifities (GHQ-12 raising global burden whereas Brief-PHQ and HADS screening for anxiety and depression), the aim of this study was to investigate the discriminating validity of the HADS, GHQ-12 and Brief-PHQ in a larger sample including the most frequent somatic diseases (cancer, cardiology, orthopaedics, endocrinology and pneumology). The analysis focused on the detection of

(1) any mental disorder and (2) anxiety as well as affective disorders. Standardized clinical interview diagnoses were the supposed criterion (gold standard) for validity of all analyses.

Methods

Sample

The study sample was recruited from 13 rehabilitation inpatient clinics in south-western Germany within a larger epidemiological study investigating the prevalence of mental disorders in chronically ill patients. The survey was based on a two step procedure. All newly admitted patients were consecutively tested using screeners (HADS, GHQ-12 and Brief-PHQ) as well as other self report measures (e.g. quality of life – SF-36) first (step 1). This happened within the first week of their stay at the clinic. In a second step, a standardized clinical interview (CIDI) was carried out by trained interviewers with randomly selected patients as to gain the diagnoses according to DSM-IV and ICD-10.

Measures

Hospital Anxiety and Depression Scale (HADS) [23]
The HADS is a self report measurement to assess anxiety and depressive symptoms, consisting of two subscales with each seven items (score 0-21). The items exclude somatic symptoms and therefore avoid symptom overlap between somatic illness and mood disorders. Good psychometric values have been found for the HADS with internal consistency being 0.80 for both subscales and retest reliability of r = 0.81-0.89 [39], [23]. Since no study proves the achievement of the subscales to be superior to the global score and the Brief-PHQ also has a total score (when analyzed dimensional), for this analysis the global 14-item total score is used.

General Health Questionnaire (GHQ-12) [24]
The GHQ-12 is a self report questionnaire for the assessment of global burden due to diseases in non-psychiatric settings. It is of use in identifying levels of severity as well as to detect psychiatric cases in epidemiological studies. It consists out of 12 items with a score range of 0-12. The psychometric values are good, with internal consistency of 0.85 and retest reliability of r = 0.73 [40], [41].

Patient Health Questionnaire (Brief-PHQ) [30]
The Brief-PHQ has been developed from the PRIME-MD System for diagnosing mental disorders. It consists of two subscales: one for the assessment of panic disorders and one for affective disorders. Additionally information about psychosocial functioning is raised. The Brief-PHQ can be used for dimensional as well as categorical diagnostics. Using it as a dimensional instrument, scores between 0 and 27 are possible. Psychometric values have been found to be good [30]. In this study the Brief-PHQ is used as a dimensional instrument as to be able to compare it to the other screeners.

Composite International Diagnostical Interview (M-CIDI) [42]
The M-CIDI is and expanded, updated and computerized version of the World Health Organization CIDI [43]. It provides reliable and efficient assessment of symptoms, syndromes and diagnoses of 48 psychiatric disorders (current, past year and lifetime prevalences). Furthermore it provides information about age of onset as well as duration and severity of symptoms. Several studies examining the psychometrics of the CIDI found satisfactory validity, with interrater reliability of kappa = 0.82-0.98 [42], [43], [44].

Data analysis

Receiver operating characteristics (ROC) analysis was used to analyse the diagnostic accuracy of the three screener HADS, GHQ-12 and Brief-PHQ in comparison to the diagnoses achieved with the standardized clinical interview CIDI.

In a ROC curve the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points of a parameter. Sensitivity is the probability that a test result will be positive when the disease is present (true positive rate, expressed as a percentage). Specifity on the other hand is the probability that a test result will be negative when the disease is not present (true negative rate, expressed as a percentage). The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups (diseased/healthy). Values range between 0.5 (no discriminatory ability) to 1.0 (perfect discriminatory ability). Each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold, in this case cut-offs of the screener. The probability, that the test is positive, when the disease is present, is called the positive predictive value (PPV). Three ROC analyses were performed. The psychiatric diagnoses from the CIDI (current 4-week prevalence) were used as case definition (criterion). The first analysis used all patients with any mental disorder as criterion. The second analysis focuses on patients with current affective disorders (major depression, dysthymia etc.) and in the third analysis patients with anxiety disorders were referred as cases (panic disorder, simple phobias etc.).

The ROC curves for each diagnostic group and the corresponding AUC were compared. According to findings that question the validity of the two HADS subscales the global score (0-42) was used. The cut-off points that minimize false positive and false negative tests will be reported as well as the corresponding sensitivity, specifity, positive and negative predictive value (PPV and NPV).

Table 1: Sample characteristics (N=204)

Sex (%) F/M	48:52
Age, mean (range) in years	49.6 (18-80)
Rates of current DSM-IV diagnose (4-week prevalence)	
Any mental disorder	28.9%
Affective disorders (single episode or recurrent major depression, dysthymia)	17.2%
Anxiety disorders (panic disorders, generalized anxiety disorders, specific phobias)	7.8%
Other (somatoform, substance related disorders)	3.9%

Table 2: Psychological strain assessed with HADS, Brief-PHQ and GHQ-12

Means (SD)	HADS	Brief-PHQ	GHQ-12
Any mental disorder N=59	21.2 (7.0)	14.5 (5.9)	8.1 (3.3)
No psychiatric diagnosis N=145	12.4 (6.5)	7.6 (4.8)	5.0 (3.7)
Affective disorder N=35	21.4 (7.1)	16.1 (5.7)	8.9 (3.0)
Anxiety disorder N=16	23.1 (5.8)	16.6 (5.1)	8.9 (2.2)

Results

Patients

A total of 204 patients participated in the two-stage examination. The mean age was 49.6 (range: 18-80), with a male/female ratio of 52% to 48% (Table 1). Somatically, 5.9% were suffering from cardiovascular diseases, 8.8% from orthopaedic diseases, 5.4% had cancer, 18.6% an endocrinologic disease and 53.4% had a pneumological disease (due to the fact, that pneumological diseases were the focus of the study). 7.8% of the patients had other somatic diseases. 28.8% of the patients received a diagnosis for a mental disorder corresponding to DSM-IV (4 weeks prevalence), 17.2% had an affective disorder and 7.8% an anxiety disorder.
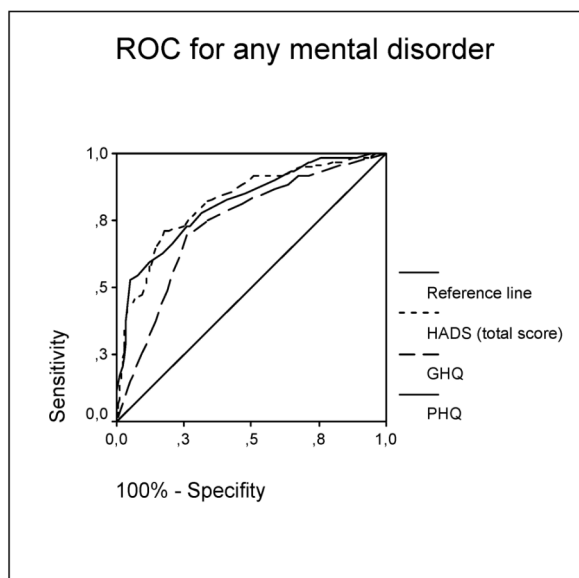
Receiver operating characteristics analyses (ROC)

For the analysis of sensitivity, specifity and positive predictive value (PPV) patients with a current psychiatric diagnosis (N=59) were divided into groups of (a) patients with affective disorders (N=35) and (b) patients with anxiety disorders (N=16).

The mean scores of HADS, Brief-PHQ and GHQ-12 were significantly higher in patients with a current psychiatric diagnosis compared to patients with no mental disorder (p<0.001; ANOVA) (Table 2). Additionally patients with affective and anxiety disorders differ significantly from patients without these disorders (p<0.000; ANOVA).

Figures 1-3 report the ability of the three screeners HADS, Brief-PHQ and GHQ-12 to identify patients with any mental disorders (Figure 1), affective disorders (Figure 2) and anxiety disorders (Figure 3) correctly. The overall accuracy (AUC) of the HADS and Brief-PHQ in detecting patients with any mental disorder differs markedly from the GHQ-12, with only the Brief-PHQ reaching significance level ($p_{PHQ/GHQ-12}$=0.038, $p_{PHQ/HADS}$=0.871, test for equalities of area).

	AUC	95% CI	p
HADS	0.820	0.760 – 0.870	0.871 (PHQ/HADS)
PHQ	0.814	0.754 – 0.856	0.038* (GHQ/PHQ)
GHQ	0.736	0.670 – 0.795	0.065 (HADS/GHQ)

* significant (p<0.05)

Figure 1: Receiver Operating Characteristics (ROC) for the detection of *any mental disorder*
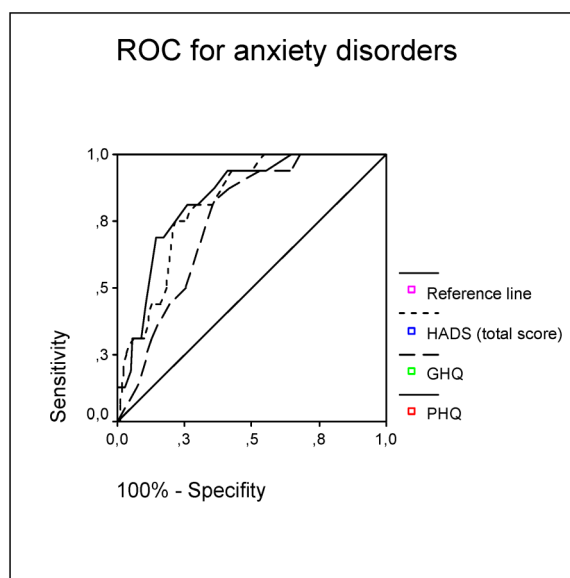
	AUC	95% CI	p
HADS	0.820	0.761 – 0.871	0.797 (PHQ/HADS)
PHQ	0.839	0.781 – 0.886	0.223 (GHQ/PHQ)
GHQ	0.753	0.688 – 0.810	0.426 (HADS/GHQ)

n.s.

Figure 3: Receiver Operating Characteristics (ROC) for *anxiety disorders*
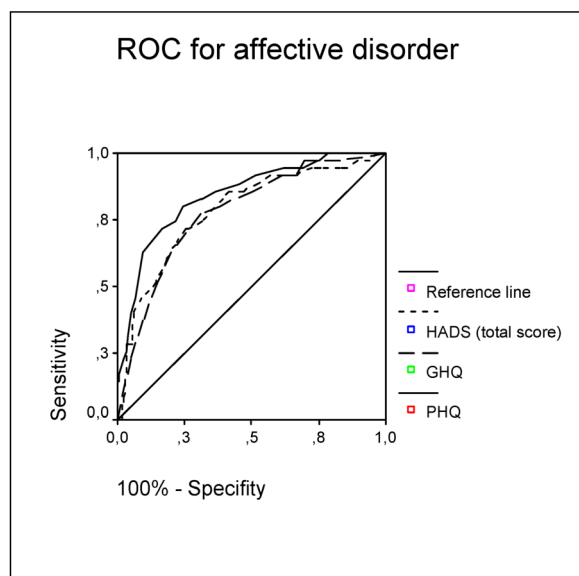
	AUC	95% CI	p
HADS	0.785	0.722 – 0.839	0.197 (PHQ/HADS)
PHQ	0.844	0.786 – 0.891	0.149 (GHQ/PHQ)
GHQ	0.779	0.716 – 0.834	0.913 (HADS/GHQ)

n.s.

Figure 2: Receiver Operating Characteristics (ROC) for *affective disorders*

Also, the ROC-analysis for the detection of affective disorders (Figure 2) shows, that the AUC of the Brief-PHQ surpasses the AUCs of the HADS and the GHQ-12, without reaching significant level though ($p_{PHQ/HADS}$=0.197, $p_{PHQ/GHQ-12}$=0.149, test for equalities of area). Corresponding results were found for the ability of the three screeners in detecting anxiety disorders. Again, the Brief-PHQ-curve exceeds the curves of the HADS and the GHQ-12, without reaching significant levels though ($p_{PHQ/HADS}$=0.797, $p_{PHQ/GHQ-12}$=0.223, test for equalities of area).

Table 3 reports the best balanced cut-offs for the three screeners along to the results considering sensitivity, specifity as well as positive predictive values (PPV) and negative predictive values (NPV). As to make sure that any mental disorder or an affective disorder is recognised with the HADS, a cut-off of 18 should be chosen (HADS-total score). For the recognition of patients with an anxiety disorder, the cut-off should be 20 considering the HADS-total score. The highest positive predictive value (PPV) was found for the HADS in recognizing any mental disorder (61.8%).

Considering the Brief-PHQ, the best cut-off points varied between 11 (recognition of affective disorders) and 15 (recognition of any mental disorder). As to identify patients with an anxiety disorder, a cut-off of 12 is best being chosen. Compared to the HADS and the GHQ-12, the Brief-PHQ has the best results in recognising any mental disorder, meaning that in 81.6% of all cases the Brief-PHQ is able to identify patients with any mental disorder correctly (PPV). Since the Brief-PHQ can either be used

Table 3: Sensitivity, specifity, positive predictive value (PPV) and misclassification

N=204	4-week prevalence N (%)	Best cut-off	Sensitivity (%)	Specifity (%)	PPV (%)	Misclassification (%)
HADS						
Any mental disorder	59 (28.9)	18	71.2	82.1	61.8	28.4
Affective disorder	35 (17.2)	18	71.4	74.6	36.8	23.2
Anxiety disorder	16 (7.8)	20	75.0	78.7	23.1	27.4
Brief-PHQ						
Any mental disorder	59 (28.9)	15	55.5	95.2	81.6	27.3
Affective disorder	35 (17.2)	11	80.0	75.7	40.6	21.0
Anxiety disorder	16 (7.8)	12	81.3	73.9	21.0	21.2
GHQ-12						
Any mental disorder	59 (28.9)	7	69.5	73.8	51.9	33.2
Affective disorder	35 (17.2)	7	77.1	69.2	34.2	24.9
Anxiety disorder	16 (7.8)	7	81.3	64.9	16.5	43.0

as categorical or as dimensional instrument, additionally its ability in identifying patients with affective and anxiety disorders were checked concerning categorical analysis. Good sensitivity and specifity values were found in recognizing major depression (sensitivity: 78%, specifity: 81%, PPV: 29%), whereas the Brief-PHQ showed to have problems to identify patients with panic disorders in this sample (sensitivity: 33%, specifity: 90%, PPV 13%).

Using the GHQ-12 for the identification of patients with any mental disorder, an affective or an anxiety disorder, a cut-off of 7 can be recommended.

Sensitivity of the screeners differ somewhat at their best cut-off points, according to diagnostic group (Table 3). Sensitivity score for recognition of any mental disorder is better in the HADS than in the GHQ-12 and the Brief-PHQ (71.2% vs. 69.5 vs. 55.5%). Nevertheless, the positive predictive value for the Brief-PHQ is clearly better than of the HADS and GHQ-12 (81.6% vs. 61.8% vs. 51.9%), meaning a higher probability of a positive test result in the Brief-PHQ than in the two other screeners, in case any mental disorder is present.

The Brief-PHQ has a higher sensitivity for recognizing affective disorders (80%) than the GHQ-12 (77.1%) which is surpassing sensitivity of the HADS (71.4%). Also the positive predictive value scores clearly higher in the Brief-PHQ than in the HADS and GHQ-12 (40.6% vs. 36.8% vs. 34.2%). In identifying anxiety disorders, the Brief-PHQ and GHQ-12 scores of sensitivity is equally 81.3%, whereas the Brief-PHQ has a somewhat higher positive predictive value than the GHQ-12 (21.0% vs. 16.5%). That means that it is more likely that the Brief-PHQ shows a positive result in patients that really do suffer from an anxiety disorder. Concerning specifity, thus the probability of recognizing patients without any disorder and without

affective disorder, the Brief-PHQ surpasses the HADS and the GHQ-12 (95.2% vs. 82.1% vs. 73.8%, resp. 75.7% vs. 74.6% vs. 69.2%). Only in identifying patients who are not suffering from an anxiety disorder, the HADS surpasses the Brief-PHQ slightly and the GHQ-12 clearly (specifity: 78.7% vs. 73.9% vs. 64.9%). Misclassification rates of the instruments vary between 21.0% (Brief-PHQ) and 42.0% (GHQ-12).

Discussion

This article describes the validation of three screeners, the HADS, Brief-PHQ and GHQ-12 for a population of chronically ill patients using ROC analysis. All patients were taken from rehabilitation inpatient clinics in the south-west of Germany. Indications of the sample were cardiology, orthopedics, oncology, endocrinology and pneumology. Patients with a pneumologic disease made out the biggest part of the sample due to study plan. As a gold standard for the ROC analyses the diagnoses of a standardized clinical interview (M-CIDI) were chosen.

The results are similar to earlier findings of our group with the HADS clearly surpassing the GHQ-12 in discriminating patients with and without mental disorders. That is the case for all of our chosen criteria, meaning the recognition of patients with (1) any mental disorder, patients with (2) an affective disorder and patients with (3) an anxiety disorder. The Brief-PHQ significantly surpasses the GHQ-12 in recognising any mental disorder and also the HADS, even in this case results do not reach significance level. Sensitivity and specifity is levelled out in a good manner for all screeners, meaning that the proportion of correctly-positive and correctly-negative assigned patients

is good. An exception is the performance of the Brief-PHQ considering recognition of any mental disorder and the performance of the GHQ-12 in the recognition of anxiety disorders.

Comparably to the study of Gräfe et al. [45] very good sensitivity and specifity values are found for the Brief-PHQ when evaluated categorically.

Best cut-offs were calculated for every screener. For the HADS, the best cut-off point is 18 for identifying any mental disorder or an affective disorder and 20 for recognizing patients with an anxiety disorder. For the GHQ-12 a cut-off of 7 was found to be best for all our chosen groups. The best cut-off points for the Brief-PHQ vary between 11 (recognition of affective disorders) and 15 (recognition of any mental disorder; 12 for the recognition of anxiety disorder) and therewith the Brief-PHQ being less consistent than the HADS and the GHQ-12 in this population. This variation of cut-offs in the Brief-PHQ makes the handling of the instruments when used for different settings complicated. This might be an important factor that needs to be looked at when choosing the Brief-PHQ.

The results of this study need to be viewed with some restrictions. First of all, the sample of this study is not homogenous, meaning that there is a strong overlap of patients with pneumologic disorders. That is due to the study plan, studying prevalence rates of mental disorders in patients with pneumologic and endocrinologic disorders. Also the chosen criteria for diagnosis of mental disorders were quite strict, excluding for instance subthreshold disorders as well as adjustment disorders. Subthreshold disorders might not be found with screeners that have been developed as categorical or assessment tools. Moreover the used screeners were analysed in a dimensional way, whereas mental diagnoses are categorical. This is an important methodological limitation of dimensionally interpreted screeners. The Brief-PHQ, which proved to be better than the HADS and GHQ-12 in this study, can also be analyzed categorical, which again is an important advantage of this instrument. Also in this study we chose to look at unspecific criteria by referring to "any mental disorder", which might explain the performance of the HADS, which performs best under genuine screening criteria. By examing the single scales, different results might have shown up. Another methodological limitation is the fact that due to the small sample size age and sex could not be controlled. Also it should be kept in mind, that the GHQ, other than the two other screeners, has not been created for the detection of mental disorders and their screening. This might explain the advantages of the HADS and the PHQ in this study. Nevertheless is this study the first to our knowledge that compared performances of the HADS, GHQ-12 and the Brief-PHQ in a sample of somatic ill patients. The results show that the Brief-PHQ as well as the HADS can be recommended as efficient instruments to detect mental disorders in somatically ill patients.

Conclusion

Taking into account the negative effects that unrecognized mental disorders have on chronical illness and the health care system, screening instruments should be used as a routine in the treatment of somatically ill patients. This way additional psychological assessment and psychological treatment (where needed) can be assigned before personal and socially negative effects become worse.

Notes

Conflicts of interest: None declared.

References

1. Frasure-Smith N, Lesperance F, Talajic M. Depression following myocardial infarction: impact on 6-month survival. J Am Med Assoc. 1995;270:1819-61.

2. Härter M, Reuter K, Weisser B, Schretzmann B, Aschenbrenner A, Bengel J. A descriptive study of psychiatric disorders and psycho-social burden in rehabilitation patients with musculoskeletal diseases. Arch Phys Med Rehab. 2002;83:461-8.

3. Parker JC, Wreight GE. Depression in arthritis and musculoskeletal disorders. In: Robertson MM, Katonal CLE, editors. Depression and Physical Illness. Chichester: John Wiley & Sons; 1997. p. 377-90.

4. Van't Spaijker A, Trijsberg RW, Duivenvoorden HJ. Psychological sequelae of cancer diagnosis: A meta-analytical review of 59 studies after 1980. Psychosom Med. 1997;59:280-93.

5. Wells KB, Golding JM, Burnam MA. Psychiatric disorder in a sample of the general population with and without chronic medical conditions. Am J Psychiat. 1988;145:976-81.

6. Carosella AM, Lachner JM, Feuerstein M. Factors associated with early discharge from a multidisciplinary work rehabilitation program for chronic low back pain. Pain. 1994;57:69-76.

7. Van Houdenhove B, Onghena P. Pain and Depression. In: Robertson MM, Katona CLE, editors. Depression and Physical Illness. Chichester: John Wiley & Sons; 1997. p. 465-97.

8. Heger S. Zur Psychosomatik des Failed-back-Syndroms: Warum Rückenschmerzen chronifizieren. Nervenarzt. 1999;70:225-32.

9. Ehlert U. Psychologie im Krankenhaus. Bern: Huber; 1998.

10. Saupe R, Diefenbacher A. Praktische Konsiliarpsychiatrie und -psychotherapie. Stuttgart: Enke; 1996.

11. Gatchel RJ, Polatin PB, Kinney RK. Predicting outcome of chronic back pain using clinical predictors of psychopathology. Health Psychol. 1995;14:415-20.

12. Hasenbring M, Marienfeld G, Kuhlendahl D, Soyaka D. Risk factors of chronicity in lumbar disc patients. Spine. 1994;19:2759-65.

13. Härter M, Woll S, Wunsch A, Bengel J, Reuter K. Recognition of psychiatric disorders in muskuloskeletal and cardiovascular rehabilitation patients. Arch Phys Med Rehab. 2004;85(7):1192-7.

14. Zabora JR. Screening procedures for psychological distress. In: Holland JC, editors. Psychooncology. New York: Oxford University Press; 1998. p. 653-62.

15. Reuter K, Woll S, Stadelmann S, Bengel J, Härter M. Erkennen und Behandeln psychischer Belastungen und Störungen in der orthopädischen Rehabilitation. Z Klin Psychol Psychiatr Psychother. 2002;50:313-27.

16. Ormel J, van den Brink W, Koeter MW, Giel R, van der Meer K, van de Willige G, Wilmink FW. Recognition, management and outcome of psychological disorders in primary care: a naturalistic follow-up study. Psychol Med. 1990;20:909-23.

17. Robbins JM, Kirmayer LJ, Cathebras P, Yaffe MJ, Dworkind M. Physician characteristics and the recognition of depression and anxiety in primary care. Med Care. 1994;32:795-812.

18. Tiemens BG, Ormel J, Simon GE. Occurrence, recognition, and outcome of psychological disorders in primary care. Am J Psychiatry. 1996;153:636-44.

19. Benjamin J, Maoz B, Shiber A, Antonovsky H, Mark M. Prevalence of psychiatric disorders in three primary-care clinics in Beersheba, Israel. Concurrent assessment by the General Health Questionnaire, general practitioners, and research diagnostic criteria. Gen Hosp Psychiatry. 1992;14:307-14.

20. Linden M, Maier W, Achberger M, Herr R, Helmchen, H, Benkert O. Ergebnisse einer Studie der Weltgesundheitsorganisation (WHO). Nervenarzt. 1996;67:205-15.

21. Pini S, Berardi D, Rucci P, Piccinelli M, Neri C, Tansella M, Ferrari G. Identification of psychiatric distress by primary care physicians. Gen Hosp Psychiatry. 1997;19:411-8.

22. Simon GE, Von Korff M. Recognition, management, and outcomes of depression in primary care. Arch Fam Med. 1995;4:99-105.

23. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiat Scand. 1983;67:361-70.

24. Goldberg D, Williams P. A user's guide to the General Health Questionnaire. London: Nfer-Nelson; 1988.

25. Goldberg DP, Oldehinkel T, Ormel J. Why GHQ threshold varies form one place to another. Psychol Med. 1998;28:915-21.

26. Jacobs KS, Bhugra D, Mann AH. The validation of the 12-item General Health Questionnaire among ethnic Indian woman living in the United Kingdom. Psychol Med. 1997;27:1215-7.

27. Pevalin DJ. Multiple applications of the GHQ-12 in a general population sample: an investigation of long-term retest effects. Soc Psychiatry Psychiatr Epidemiol. 2000;35:508-12.

28. Schmitz N, Kruse J, Heckrath C, Alberti L, Tress W. Diagnosing mental disorders in primary care: The General Health Questionnaire (GHQ) and teh Symptom Check List (SCL-90-R) as screening instruments. Soc Psychiatry Psychiatr Epidemiol. 1990;34:360-6.

29. Üstün TB, Sartorius N. Mental Illness in General Health Care. Chichester: John Wiley & Sons; 1995.

30. Spitzer RL, Kroenke K, Williams JBW, the Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of Prime-MD. JAMA. 1999;282(18):1737-44.

31. Rizzo R, Piccinelli M, Mazzi MA, Bellantuono C, Tansella M. The personal Health Questionnaire: A new screening instrument für detection of ICD-10 depressive disorders in primary care. Psychol Med. 2000;30:831-40.

32. Reuter K, Härter M. Screening for mental disorders in cancer patients - discriminant validity of HADS and GHQ-12 assessed by standardized clinical interview. Int J Methods Psychiatr Res. 2002;10:86-96.

33. Härter M, Reuter K, Groß-Hardt K, Bengel J. Screening of anxiety, depressive and somatoform disorders in rehabilitation - validity of HADS and GHQ-12 in patients with musculoskeletal disease. Disabil Rehabil. 2001;23:737-44.

34. Härter M, Woll S, Wunsch A, Bengel J, Reuter K. Screening for mental disorders in cancer, cardiovascular and musculoskeletal diseases. Comparison of HADS and GHQ-12. Soc Psychiatry Psychiatr Epidemiol. 2005;1:1-7.

35. Wunsch A, Rundel M, Härter M, Bengel J. Screening psychischer Störungen in der kardiologischen und orthopädischen Rehabilitation. Rehabilitation. 2002;41:375-81.

36. Löwe B, Grafe K, Zipfel S, Witte S, Loerch B, Herzog W. Diagnosing ICD-10 depressive episodes: superior criterion validity of the Patient Health Questionnaire. Psychother Psychosom. 2004;73(6):386-90.

37. Löwe B, Spitzer RL, Grafe K, Kroenke K, Quenter A, Zipfel S, Buchholz C, Witte S, Herzog W. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. J Affect Disor. 2004;78(2):131-40.

38. Löwe B, Grafe K, Zipfel S, Spitzer RL, Herrmann-Lingen C, Witte S, Herzog W. Detecting panic disorder in medical and psychosomatic outpatients: comparative validation of the Hospital Anxiety and Depression Scale, the Patient Health Questionnaire, a screening question, and physicians' diagnosis. J Psychosom Res. 2003;55(6):515-9.

39. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale - A review of validation data and clinical results. J Psychosom Res. 1997;42:7-41.

40. Hardy GE, Shapiro DA. Validation of the General Health Questionnaire-12 using a sample of employees from England's Health Care Service. Psychol Assess. 1999;11(2):159-65.

41. Mccabe CJ, Thomas KJ, Brazier JE, Coleman P. Measuring the mental health status of a population: a comparison of the GHQ-1 and the SF-36 (MHI-5). Br J Psychiatry. 1996;169:517-21.

42. Wittchen HU, Weigel A, Pfister H. DIA-X Interview. Frankfurt: Swets Test Services Deutschland; 1996.

43. Wittchen HU, Lachner G, Wunderlich U, Pfister H. Test retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). Soc Psychiatry Psychiatr Epidemiol. 1998;33:568-78.

44. The Esemed/Mhedea 2000. Prevalence of mental disorders in Europe: results from the European study of the epidemiology of mental disorders (ESEMeD) project. Acta Psychiatr Scand. 2004;109(Suppl 420):21-7.

45. Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". Diagnostica. 2004;50(4):171-81.

Corresponding author:

Daniela Hahn
University of Northampton, School for Social Sciences, Department for Psychology, Boughton Green Rd, Northampton NN2 7AL, UK
daniela.hahn@northampton.ac.uk