Future Academy®
Offering excellence in Social and Behavioural Sciences

The European Proceedings of
Social & Behavioural Sciences
EpSBS

eISSN: 2357-1330

icH&Hpsy 2016 : 2nd International Conference on Health and Health Psychology

Reliability and Validity of the PHQ-9 in Portuguese Women with Breast Cancer

Ana Torres[a], Sara Monteiro[b]*, Anabela Pereira[c], Emília Albuquerque[d]

* Corresponding author: Sara Monteiro, smonteiro@ua.pt

[a] *Assistant Professor Dr, University of Aveiro, ESEnfCVPOA, Cintesis, Aveiro, Portugal, anatorres@ua.pt*
[b] *Assistant Professor Dr, University of Aveiro, Cintesis, Aveiro, Portugal,smonteiro@ua.pt*
[c] *Associate Professor Dr, University of Aveiro,Cintesis, Portugal*
[d] *Oncological Institute of Portugal IPO,Portugal*

Abstract

http://dx.doi.org/10.15405/epsbs.2016.07.02.39

Background: Depression is often overlooked in oncology practice. Depressive module PHQ-9 was not studied in Portuguese cancer patients, despite its brevity and comparable psychometric proprieties to other depression scales. Which are the psychometric properties of the Portuguese version of the PHQ-9 for use with women with breast cancer? Reliability, sensitivity to change and the construct validity will be studied. The purpose of this study is to assess the psychometric properties of the Portuguese version of the PHQ-9 for use with women with breast cancer. This study intends to evaluate the reliability, the sensitivity to change and the construct validity of the PHQ-9 in a Portuguese Breast Cancer women sample. **Methods:** It was reproduced a principal component analysis and explored the reliability and validity of the questionnaire. The validation used a sample of 63 Portuguese women with breast cancer. A test-retest was conducted in 45 women, after 8 weeks. Construct validity was analysed.
Results: PHQ-9 scores ranged from 0-27, with a mean score of 4,92 (SD= 4,63). The scale presented adequate internal consistency (α = .86) and test retest reliability (ICC= .87). It also presented good construct validity, as overall scores and severity levels were strongly associated with functional and symptoms subscales. The principal component analysis explains 48.42% of the variance. **Conclusions:** The validation process of the Portuguese PHQ-9 version shows metric properties similar to those in international studies, suggesting that it measures the same constructs, in the same way, as the original version. Data provided evidence for the validity of the PHQ-9 as a brief measure of depression severity in Portuguese women with breast cancer.

Keywords: breast cancer; PHQ-9; depressive symptoms; psychometric.

1. Introduction

Depression is the greatest cause of incapacity in the world according to World Health Organization (Murray& Lopez, 1996). Given the high prevalence and its consequences to the general population,

this mental illness is considered to be one of the priority goals of the National Health Plan, which gives rise to the urgent need for an integrative assessment and for a management approach in the scope of mental illness (Ministério da Saúde [Health Ministry], 2004). The lifetime prevalence rates for major depressive disorder in the general population are high, about 16.6% (Kessler, Chiu, Demler, & Walters, 2005), and prevalence rates in cancer patients are even higher (Hinzet al., 2010; Van't Spijker, Trjsburg, & Duivendoorden, 1997), ranging from 0% to 38% for major depression, and 0% to 58% for depression spectrum syndromes (Massie, 2004). According to Massie (2004), breast cancer is one of the cancer types that is highly associated with depression (1.5%-46%). Depression has an impact on quality of life and on the satisfaction with, and participation in medical treatment (Bui, Ostir, Kuo, Freeman & Goodwin, 2005; Skarstein, Aass, Fossa, Skovlund, & Dahl, 2000). Despite these facts, depression is often overlooked (Fallowfield, Ratcliffe, Jenkins, & Saul, 2001) and untreated in oncology practice, especially when it is most needed (Söllner et al., 2001; Passik et al., 1998).

The use of case-finding instruments for the detection of depression is strongly suggested (Mitchell, 2007). There are a number of instruments to detect depression symptoms, but the Patient Health Questionnaire's depressive module (PHQ-9) is half the length of many of the other depression scales and has comparable sensitivity and specificity. The Patient Health Questionnaire (PHQ; Kroenke, Spitzer, & Williams, 2001) is a self-report instrument developed to screen for psychiatric disorders in primary care and assess eight specific disorders based on DSM-IV (APA, 2002) criteria. It was validated in primary care patients and it has been used with cancer patients (Ell et al., 2007; Spitzer, Kroenke, & Williams, 1999). The nine items for depression (PHQ-9) and the 22 items for anxiety disorders may easily be translated into oncology settings, unlike the modules for somatoform disorder and eating disorders (Pirl, 2010). The diagnostic validity of PHQ has been established in studies involving a large number of patients (Spitzer et al., 1999; Spitzer, Williams, Kroenke, Hornyak, & McMurray, 2000).

The PHQ-9 consists of a depression module with nine items, using a 4-point response scale, ranging from "0" (not at all) to "3" (nearly every day) to report each of the 9 DSM criteria. It can be scored both categorically, to provide likely diagnoses, and continuously, to give level of symptoms. To meet the criteria for a probable major depression diagnosis patients must state having the two items for low mood and anhedonia, and report having at least five other symptoms "at least half of the time" or more, which is consistent with DSM-IV criteria (Pirl, 2010). As a severity measure, the PHQ-9 score can range from 0 to 27, with high scores being related to greater severity of depression (Kroenke et al., 2001). It also includes a final item asking patients who reported any problems: "How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?". Regarding psychometric characteristics, the internal consistency is excellent, with a Chronbach's alpha of .89 in the PHQ Primary Care Study and .86 in the PHQ Obstetric-Gynaecological Study (Kroenke et al., 2001). Different studies (Cameron, Crawford, Lawton, & Reid, 2008; Lee, Schulberg, Raue, & Kroenke, 2007) found good internal consistency reliability. The test-retest reliability has been reported with a Correlation Coefficient from .71 (Omoro, Fann, Weymuller, Macharia, & Yueh, 2006) to .92 (Pinto-Meza, Serrano-Blanco, Penarrubia, Blanco, & Haro, 2005). Convergent validities were also found, from .56 (Lotrakul, Sumrithe, & Saipanish, 2008) to .74

http://dx.doi.org/10.15405/epsbs.2016.07.02.39
eISSN: 2357-1330 / Corresponding Author: Sara Monteiro
Selection and peer-review under responsibility of the Organizing Committee of the conference

(Changsu et al., 2008). A meta-analysis of Gilbody, Richards, Brealey and Hewitt (2007) found a high sensitivity (.80) and specificity (.92). Moreover, a systematic review of Wittkampf, Naeije, Schene, Huyser and van Weert (2007) found a PHQ-9 sensitivity of .77 and a specificity of .94. We found studies with sensibilities ranging from .5 (Carballeira et al., 2007) to 1 (Osório, Mendes, Crippa, & Loureiro, 2009), and with specificities ranging from to .76 (Carballeira et al., 2007) to .98 (Yeung et al., 2008; Osório et al., 2009). More recently in a systematic review, Kroenke, Spitzer, Williams and Löwe (2010) reaffirm that PHQ-9 has a good sensitivity and specificity for detecting depressive disorders, as well as a good sensitivity to change, which lead to the conclusion that this is a brief, well-validated measure for detecting and monitoring depression. In this study the authors further restated the PHQ-9 cut-off points of 5, 10 and 15 for mild, moderate and severe symptom levels, respectively.

Several PHQ-9 versions were validated in different samples and in diverse languages and cultural contexts, demonstrating proprieties that confirm it as a reliable scale to assess depressive symptoms, namely in: primary health care (Cameron et al., 2008; Diez-Quevedo, Rangil, Sanchez-Planell, Kroenke, & Spitzer, 2001; Liu et al., 2011; Osório et al., 2009; Yu, Tam, Wong, Lam, & Stewart, 2012); students (Adewuya, Ola, & Afolabi, 2006; Granillo, 2012); elderly population (Changsu et al., 2008). Moreover it has been adapted to different languages and cultures, such as Brazilian (Osório et al., 2009), Thai (Lotrakul et al., 2008), Greek (Karekla, Pilipenko, & Feldman, 2012), Saudi Arabian (Becker, Zaid, & Faris, 2002), Taiwanese (Liu et al., 2011), Chinese (Yeung et al., 2008), Korean (Changsu et al., 2008), Spanish (Diez-Quevedo et al., 2001), Mexican (Donlan & Lee, 2010), French (Carballeira et al., 2007), Nigerian (Adewuya et al., 2006), Turkish (Güleç, Güleç, Simsek, Turhan, & Sünbül, 2012) and Japanese (Yu et al., 2012). The PHQ-9 has been reported to have a single factor in several previous studies (Cameron, Crawford, Lawton, & Reid, 2008; Dum, Pickren, Sobell, & Sobell, 2008; Graves & Bombardier, 2008; Hansson, Chotai, Nordstöm, & Bodlund, 2009; Yu et al., 2012), including a study involving more than 5000 primary care patients (Huang, Chung, Kroenke, Delucchi, & Spitzer, 2006). Exceptions are typical among participants with Spinal Cord Injury, in which the most of the studies have suggested a 2-factor structure, differentiating somatic and non somatic symptoms (Krause, Reed & McArdle, 2010; Richardson, & Richards, 2008). Kalpakjian et al. (2009), after comparison of 1- and 2-factor solution, concluded that the 1-factor solution represented a better fit, because of sex congruence.

2. Problem Statement

Despite its strong psychometric properties and widespread use in primary care, there is scarcity of information when used in cancer patients, as it was concluded in the major recent review of screening questionnaires for psychological distress in cancer patients (Vodermaier, Linden, & Siu, 2009). The few studies reported, however, good psychometric properties of this scale with this population. For instance, even the original validation encloses obstetrics and gynaecology patients and found to have strong psychometric proprieties (Kroenke et al., 2001). The PHQ-9 also demonstrated adequate reliability as well as concurrent and divergent validity in a study of head and neck cancer patients (Omoro et al., 2006) and in a study that used a touch screen computerized version of the questionnaire with 342 cancer patients (Fann et al., 2009). A more recent study (Johns et al., 2013) found an internal

consistency of .81 and good correlations with other measures of depressive symptoms, which lead them to consider it as a reliable and valid depressive measure in adults with cancer, as well as a useful measure for monitoring depression therapy and outcomes in this population.

3. Research Questions

Which are the psychometric properties of the Portuguese version of the PHQ-9 for use with women with breast cancer? Reliability, sensitivity to change and the construct validity will be studied.

4. Purpose of the Study

The purpose of this study is to assess the psychometric properties of the Portuguese version of the PHQ-9 for use with women with breast cancer. This study intends to evaluate the reliability, the sensitivity to change and the construct validity of the PHQ-9 in a Portuguese Breast Cancer women sample.

5. Research Methods

5.1. Sample

The participants of the study, 63 women with breast cancer, were recruited from the Reach to Recovery Movement of the Portuguese Cancer League (Mean age= 60.46, SD=10.88, Min=30, Max=85). To be eligible to participate, patients had to be adults (older than 18 years) and have a breast cancer diagnosis history. Patients were excluded if they had psychotic disorder, neurological disease or substance abuse that might compromise their ability to complete the questionnaires, or if they declined to participate voluntarily after informed consent. Also, we excluded those patients who did not answer more than 50% of the assigned questionnaires. The majority of the sample was married (81%), attended tertiary education (31.7%) and retired (85.7%). Regarding the clinical characteristics, the sample is composed by survivors (88.9%) and patients in the treatment phase (11.1%), given that the majority: self-detected their cancer (39.7%); realized mastectomy (73%); did axillary lymph node dissection (58.7%), chemotherapy (68.3%) and did no radiotherapy (54%). Just 11.1% had cancer recurrence. All demographic and clinical variables are detailed in Table 1. A total 45 women participated in the test-retest study, after 8 weeks. Moreover, 2 women did not answer completely to the questionnaires used for study's the construct validity.The sample of participants consisted of 33 older adults, aged from 65 to 94 years old (M=82.18, SD=7.85), participated in 2 groups, group with intervention (GI, n=16) and a waiting group (GC, n=17).

Table 1. Sociodemographic and clinical variables

Sociodemographic variables	n=63
Age	
Mean; SD	60.46; 10.88
Min-máx	30-85
Marital Status	n(%)
Single	3 (4.8)
Married	51 (81)
Divorced	3 (4.8)
Widow	6 (9.5)

http://dx.doi.org/10.15405/epsbs.2016.07.02.39
eISSN: 2357-1330 / Corresponding Author: Sara Monteiro
Selection and peer-review under responsibility of the Organizing Committee of the conference

Clinical variables		n (%)
Detection of cancer type		
Self-examination		25 (39.7)
Other		12 (19)
Screening tests ordered by the doctor		10 (15.9)
Screening by the Portuguese Cancer League	Examination/palpation performed	8 (12.7)
by the physician in	consultation routine	10 (15.9)
Type of surgery performed		
Mastectomy		46 (73)
Lumpectomy		6 (9.5)
Quadrantectomy		5 (7.9)
Double mastectomy		3 (4.8)
Quadrantectomy + Mastectomy		2 (3.2)
Type of treatment		
Radiotherapy		29 (46)
Chemotherapy		43 (68.3)
Hormone therapy		31 (49.2)
Axillary dissection		37 (58.7)

5.2. Assessment tools

The questionnaire covered the socio-demographic information, such as age, gender, marital status and study cycle, as well as, the depressive symptoms and the quality of life. Cancer-specific variables included treatment status (on or off treatment), way of cancer detection and type of treatments received.

Depressive symptoms were assessed by two measures, the PHQ-9 (in study on this work) and the Hospital Anxiety and Depression Scale [HADS] (Zigmond & Snaith, 2003). HADS consists of two subscales, one assessing anxiety, and one measuring depression, both with seven items, which are scored separately. Each item is answered by the patient on a 4-point (0 – 3) response scale, thus, the possible scores range from 0 to 21 for anxiety and 0 to 21 for depression. The HADS manual indicates that a score between 0 and 7 is ‘‘normal’’, between 8 and 10 ‘‘mild’’, between 11 and 14 ‘‘moderate’’, and between 15 and 21 ‘‘severe’’. Zygmond and Snaith (2003) posit a score of 11 or higher indicating probable presence (‘‘caseness’’) of a mood disorder and a score of 8 – 10 being just suggestive of the presence of the respective state. The original version, as does the Portuguese version (Pais-Ribeiro et al., 2007), presents adequate psychometric characteristics.

The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core-30 [EORTC QLQ-C30] (Aaronson et al., 1993) was used to assess the QOL. It is a 30-item questionnaire, 24 of the items constitute nine multi-item scales, and six items are single-item symptom measures. The scales are constructed by sum of the scores of the items. Multi-item subscales and single items intend to reflect the multidimensionality of the QOL construct, namely: five functional subscales (physical, role, cognitive, emotional, and social); a global health/QOL subscale; three symptom subscales (fatigue, pain, and nausea/vomiting); and single items for the assessment of additional symptoms commonly reported by cancer patients (dyspnoea, appetite loss, sleep disturbance, constipation, and diarrhoea); one more item related to the perceived financial impact of cancer and cancer treatments. All the items are scored on 4-point Likert type scale, ranging from 1 “not at all” to 4 “very much”, except for the two items of the global health/quality of life subscale, that uses a modified 7-point linear analogue scale. A high score for functional scales and global health status/QOL represents a high/healthy level of functioning and quality of life, in opposition to symptom subscales

and single items. The original version, as does the Portuguese version (Pais-Ribeiro, Pinto, & Santos, 2008), of EORTC QLQ-C30 presents adequate psychometric characteristics.

5.3. Procedure

All statistical analyses were conducted with SPSS, version 19.0. We adhered to the classical test theory to examine the reliability and validity of the PHQ-9, guided for Carretero-Dios and Pérez (2007), and similarly to the Kenyan validation by Omoro et al. (2006). Reliability was established by assuring both internal consistency and testretest reliability. The validity was evaluated by the construct validity. We hypothesized that higher PHQ-9 scores were related to lower results in the functional scales of quality of life and higher symptoms (evaluated by HADS and QLQ C-30 symptoms subscales). The cut-off points suggested by the original authors were the ones adopted, to identify minimal (0-4), mild (5-9), moderate (10-14), moderately severe (15-19), and severe (20-27) depression. Construct validity of the PHQ-9 as a measure of depression severity was assessed by examining functional sub-scales and symptoms sub-scales (evaluated by HADS, QLQ C-30 and supl. BR-23), over the 5 PHQ-9 intervals. To test the differences of the 5 groups of severity on the quality of life, we opted for a Jonckheere-Terpstra test, which is similar to the Kruskal-Wallis test, but incorporates information about whether the order of the groups is meaningful. In parallel, in order to evaluate the differences of the 5 groups of severity of PHQ-9 on the 4 depression levels evaluated by the HADS, we chose to use the Pearson's chi-square test. We used the 4 groups of severity of depression: normal (0 to 7); mild (8 to 10); moderate (11 to 14) and severe (15 to 21), suggested by the HADS authors.

Adaptation to the Portuguese language

The translation used a two-person English-Portuguese translation, a two-person retroversion, and a discussion group to achieve full consensus for the lexical and cultural equivalence. The content validity analysis was performed by two psychologists. After lexical equivalence and content validity were defined, a cognitive debriefing analysis was performed with subjects from the lower educational levels and from the oldest groups of potential patients. After the final version was defined, we chose a face format for the questionnaire identical to theoriginal.

6. Findings

6.1. Symptomatology results

PHQ-9 scores ranged from 0-27, with a mean score of 4.92 (SD= 4.63). The majority (53.1%) presents scores, which indicate minimal depression severity symptoms, 34.4% present mild severity, 9.4% moderate, and 1.6% severe.

6.2 Reliability

Chronbach's alpha was performed to identify the internal consistency of the scale. A good Chronbach's alpha of .86 was found. We also observed that none of the items correlated less than .52 with the total score (above the required minimum of .30), and none of the items would affect reliability if they were deleted (see Table 2).

http://dx.doi.org/10.15405/epsbs.2016.07.02.39
eISSN: 2357-1330 / Corresponding Author: Sara Monteiro
Selection and peer-review under responsibility of the Organizing Committee of the conference

Table 2. Reliability analysis

Items	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1	.52	.85
2	.60	.84
3	.53	.85
4	.75	.82
5	.55	.85
6	.55	.84
7	.65	.86
8	.56	.84
9	.62	.84
alpha		.86

6.3 Test-retest

In order to inspect test-retest correlation, the instruments were administered twice, with an interval of 8 weeks, in 45 breast cancer women. The intra-class correlation coefficient was excellent (ICC= .87).

6.4 Construct validity

Examining the PHQ-9 as a continuous variable, its correlation with the sub-scale of depression of HADS was .46 ($p<.001$). It also correlates significantly with sub-scale of anxiety of HADS, .42 ($p<.001$). This data can indicate that PHQ-9 is also sensitive to anxiety states. PHQ-9 also presented significant correlations with some sub-scales of EORTC QLQ-C30 and Supl. BR-23, namely negative correlations with: role functioning ($rS= -.29$, $p<.05$), emotional functioning ($rS= -.51$, $p<.001$), cognitive functioning ($rS= -.31$, $p<.05$), sexual enjoyment ($rS= -.47$, $p<.05$), and future perspective ($rS= -.36$, $p<.05$). Significant positive correlations were also obtained in the following sub-scales: fatigue ($rS= .52$, $p<.001$), constipation ($rS= .29$, $p<.05$), systemic therapy side effects ($rS= .33$, $p<.05$), breast symptoms ($rS= .53$, $p<.001$) and arm symptoms ($rS= .35$, $p<.05$). The associations between increasing PHQ-9 depression severity scores and EORTC QLQ C-30, BR-23 and depression and anxiety severity of HADS are shown in the Table 3, 4 and 5.

Table 3. Relationship between PHQ-9 depression severity levels and QLQ C-30 functional scales

Level of Depression Severity, PHQ-9 Score	EORTC QLQ C-30 Scores					
	Global health status	Functional scales				
		Physical functioning	Role functioning	Emotional functioning	Cognitive functioning	Social functioning
		(n, M, Md)				
Minimal 0-4	61, 70.49, 75	60, 83.44, 86.67	57, 88.01, 100	61, 76.09, 83.33	59, 83.05, 83.33	61, 92.62, 100
Mild 5-9	22, 68.94, 70.84	21, 82.86, 86.67	20, 89.17, 91.67	22, 70.08, 75	21, 78.57, 83.33	22, 92.42, 100
Moderate 10-14	6, 54.17, 54.17	6, 80, 83.33	5, 66.67, 66.67	6, 59.72, 58.33	5, 66.67, 66.67	6, 77.76, 91.67
z	-1.46	-.60	-2.03	-3.47	-2.76	-1.69
p	.15	.55	.04	.00	.00	.09

n= number of participants, M= Mean, Md= Median, z = z-core of Jonckheere-Terpstra test, p= significance level

In Table 3 it is possible to observe that there are significant differences between the severity of groups regarding role functioning, emotional functioning and cognitive functioning. The negative z-value should also be noted, given that it indicates a trend of decreasing medians with the increment of depression severity. Several significant differences are also observed in the severity groups regarding symptoms, namely, fatigue, dyspnoea and constipation, as shown in Table 4. In this case, the z-value is positive, indicating a trend towards ascending medians with the increment of depression severity.

Table 4. Relationship between PHQ-9 depression severity levels and QLQ C-30 symptoms scales

Level of Depression Severity, PHQ-9 Score	EORTC QLQ C-30 Scores								
	Symptom scales / items								
	Fatigue	Nausea and Vomiting	Pain	Dyspnoea	Insomnia	Appetite loss	Constipation	Diarrhoea	Finantial difficulties
	(n, M, Md)								
Minimal 0-4	60, 22.40, 22.22	59, 3.11, 0	59, 16.38, 16.67	60, 3.89, 0	60, 28.89, 33.33	60, 10.56, 0	61, 16.39, 0	61, 1.64, 0	61, 11.48, 0
Mild 5-9	21, 29.10, 22.22	21, 0, 0	20, 17.50, 16.67	21, 4.76, 0	21, 28.57, 33.33	21, 4.76, 0	22, 13.63, 0	22, 0, 0	22, 12.12, 0
Moderate 10-14	6, 38.89, 44.44	5, 30, 0	6, 30.56, 25	6, 16.67, 16.67	6, 55.56, 66.67	6, 44.44, 33.33	6, 55.56, 50	6, 0, 0	6, 33.33, 16.67
z	3.68	1.02	1.5	2.80	1.72	1.91	2.40	1.57	1.66
p	.00	.31	.12	.01	.09	.06	.02	.12	.10

n= number of participants, M= Mean, Md= Median, z = z-core of Jonckheere-Terpstra test, p= significance level

The same pattern of significant differences between the groups' severity is observed for quality of life, by the general QLQ C-30 scale, and is also observed in the breast cancer module (BR-23), as presented in Table 5. Only one functional scale (future perspective) and one symptom scale (upset by air loss) did not present significant differences for severity groups. The sign of z-value is the same as that for functional and symptoms scales of the QLQ C-30 measure, indicating a trend toward ascending medians with the increment of severity for functional scales and a trend of decreasing medians for symptoms scales. The relationship between depression levels of PHQ-9 and levels of HADS shows that PHQ-9 groups differ significantly on depression levels of HADS, but not on the anxiety levels of HADS (Table 6).

http://dx.doi.org/10.15405/epsbs.2016.07.02.39
eISSN: 2357-1330 / Corresponding Author: Sara Monteiro
Selection and peer-review under responsibility of the Organizing Committee of the conference

Table 5. Relationship between PHQ-9 depression severity levels and BR-23 scales

Level of Depression Severity, PHQ-9 Score	EORTC BR-23 Scores							
	Functional scales				Symptom scales / items			
	Body image	Sexual functioning	Sexual enjoyment	Future perspective	Systemic therapy side effects	Breast symptoms	Arm symptoms	Upset by air loss
	(n, M, Md)				(n, M, Md)			
Minimal 0-4	60, 82.92, 91.67	55, 20.30, 16.67	25, 46.67, 33.33	61, 57.92, 66.67	58, 15.35, 9.52	61, 21.18, 16.67	59, 20.53, 11.11	12, 30.56, 33.33
Mild 5-9	22, 80.68, 91.67	19, 16.67, 0	8, 45.83, 33.33	22, 46.97, 33.33	21, 17.91, 14.29	22, 26.14, 25	21, 26.98, 22.22	4, 8.33, 0
Moderate 10-14	5, 61.67, 58.33	6, 11.11, 8.33	2, 16.67, 16.67	6, 33.33, 0	4, 27.38, 33.33	6, 37.5, 37.5	6, 33.33, 16.67	1, 66.67, 66.67
z	-2.16	-2.12	-1.65	-2.51	2.24	3.82	2.02	-0.5
p	.03	.03	.09	.01	.03	.00	.02	.56

n= number of participants, M= Mean, Md= Median, z = z-core of Jonckheere-Terpstra test, p= significance level

Table 6. Relationship between PHQ-9 depression severity levels and HADS severity levels

Level of Depression Severity, PHQ-9 Score	HADS levels	
	Anxiety	Depression
	(n, Md, Mo)	
Minimal 0-4	33, 0, 0	33, 0, 0
Mild 5-9	22, 0, 0	21, 0, 0
Moderate 10-14	6, 1.5, 2	5, 1, 0[a]
$\chi 2$	11.86	16.77
Df	6	4
p	.07	.00

n= number of participants, Md= Median , Mo= Mode, $\chi 2$= Pearson chi-square test, p= significance level
a. Multiple modes exist. The smallest value is shown.

6.5 *Factorial validity*

It was performed exploratory factor analysis to identify the number of factors and the magnitude of the factorial loading, using a principal component analysis with Kaiser Normalization for the total patients (n= 63). We found a single factor solution, which explains 48.42% of the variance, with all the items presenting loadings above .40. The factor analysis results provide further information about the proprieties of the PHQ-9. The determinant of the correlation matrix is .01 (greater than necessary

value, .0001). The observation of the correlation matrix also leads us to conclude that all the items correlate fairly well with all other, and no correlation coefficients are particularly large. The value of KMO for multiple variables, as a measure of sampling adequacy, was very good (.83), therefore we should be confident that the factor analysis is appropriate for these data. Bartlett's test of sphericity was significant ($p<.001$), which indicates that Rmatrix is not an identity matrix, thus, the sample is adequate allowing for a factorial analysis.

7. Conclusions

In this study we completed a translation, test of reliability, construct validation and factorial validity of the Portuguese PHQ-9 depression scale in a breast cancer sample. Data provided evidence for the validity of the PHQ-9 as a brief measure of depression severity in Portuguese women with breast cancer. A first principal component analysis results on a single factor solution explaining 48.42% of the variance. The majority of studies of the factor analysis of PHQ-9 holds a single factor solution (Cameron et al., 2008; Dum et al., 2008; Graves & Bombardier, 2008; Hansson et al., 2009; Huang et al., 2006; Kalpakjian et al., 2009; Yu et al., 2012), with explained variances from 38.9% (Huang et al., 2006) to 59.57% (Dum et al., 2008). A onedimensional solution would be suggested, theoretically, by the DSM-IV. Exceptions reported a two factor solution (Krause et al., 2010; Richardson, & Richards, 2008). Our results lend support to the notion of a single-factor structure for the symptoms assessed by the PHQ-9.

Reliability of .86 is exactly the same value obtained in the original study, which includes an obstetrics-gynaecology cancer sample (Kroenke et al., 2001). This value is argued as excellent, because when the test is used for diagnostic or classification purposes, the minimum value advisable should be around .80 (Nunnally & Bernstein, 1995). This measure presents good test-retest reliability, with a correlation of .84. This correlation is better than other versions used with cancer patients, for instance .71 obtained by Omoro et al. (2006), but lower than versions used with other samples (Adewuya et al., 2006; Liu et al., 2011; Pinto-Meza et al., 2005).

The construct validity was established by the strong association between PHQ-9 scores, HADS and Quality of life scores. The results demonstrate that higher PHQ-9 scores are related to lower functional scales of quality of life, and increased symptoms.

Similarly to original validation (Kroenke et al., 2001), we verified a strong association between increasing PHQ-9's depression severity scores and the worsening of functional scales, evaluated by QLQ C-30 and BR-23. We also obtained significant decreases for the scales that previous studies have shown should be most strongly related to depression, i.e. mental health (emotional functioning, cognitive functioning, future perspective) and role functioning. The association with symptoms' sub-scales, especially with the depression sub-scale of HADS, was also expected and showed the extent to which both instruments measure the same construct.

Some limitations of the present study should be pointed out. The size and the composition of the sample, which consisted exclusively of women, prevent the results from being generalized to other samples or to the general population. Accordingly, future studies involving different and larger samples

are still necessary. Samples with higher levels of depression severity are still necessary too. The data are limited to self report.

It would be interesting to have external data that could be used, as a structured interview, to determine the sensibility and specificity of the measure. To perform a confirmatory factor analysis in addition to exploratory factor analysis, with a Portuguese larger sample, would be a good suggestion for future studies.

To our knowledge, this is the first study to investigate the PHQ-9 in Portuguese cancer patients. Other studies with cancer patients are encouraged. As main implication to the research and clinical practice, this study contributes with the adaptation and validation of an instrument that has characteristics (namely brevity, adequate psychometric characteristics and diagnostic potential) that are useful for the assessment and emotional screening of cancer patients.

Acknowledgments

This study was developed under the scope of a PhD fellowship granted by Fundação para a Ciência e Tecnologia [Foundation for Science and Technology of Portugal] (Reference Number: SFRH / BD / 44296 / 2008). The authors would like to thank Reach to Recovery Movement of Portuguese Cancer League [Movimento Vencer e Viver da Liga Portuguesa contra o Cancro] for allowing the access to breast cancer patients. The authors would also like to thank to Raquel V. Oliveira (psychologist and native speaker) for proof reading the paper prior to its submission.

References

Aaronson, N., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N., …Takeda, F. (1993). The European Organization for Research and Treatment of Cancer QLQ – C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 85(5), 365-376. doi:10.1093/jnci/85.5.365

Adewuya, O. A., Ola, B. A., & Afolabi, O. O. (2006).Validity of the Patient Health Questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. *Journal of Affective Disorders*, 96, 89-93. doi:10.1016/j.jad.2006.05.021

American Psychiatric Association (2000).*Manual de diagnóstico e estatística das perturbações mentais*. Lisboa: Climepsi Editores

Becker, S., Zaid, K. A., & Faris, E. A. F. (2002). Screening for somatization and depression in Saudi Arabia: A validation study of the PHQ in primary care. *International Journal of Psychiatry in Medicine*, 32(3), 271-283.doi:10.2190/XTDD-8L18-P9E0-JYRV

Bui, Q. U., Ostir, G., Kuo, Y. F., Freeman, J., & Goodwin, J. (2005). Relationship of depression to patient satisfaction: findings from the barriers to breast cancer study. *Breast Cancer Research and Treatment*, 89(1), 23-28.doi:10.1007/s10549-004-1005-9

Cameron, I. M., Crawford, J. R., Lawton, K., & Reid, I. C. (2008).Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *British Journal of General Practice, 58*(546), 32–36.doi:10.3399/bjgp08X263794

Carballeira, Y., Dumont, P., Borgacci, S., Rentsch, D., Tonnac, N., Archinard, M. & Andreoli, A. (2007).Criterion validity of the French version of Patient Health Questionnaire (PHQ) in a hospital department of internal medicine. *Psychology and Psychotherapy: Theory, Research and Practice, 80*(1), 69-77. doi:10.1348/147608306X103641

Cameron, I. M., Crawford, J. R., Lawton, K., & Reid, I.C. (2008). Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *British Journal of General Practice*, 58, 32-36. doi: 10.3399/bjgp08X263794

Carretero-Dios, H., & Perez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology, 7*(3), 863-882.

Changsu, C., Jo, S. A., Kwak, J., Pae, C., Steffens, D., Jo, I., & Park, M. H. (2008). Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Asian Geriatric study. *Comprehensive Psychiatry, 49*(2), 218-223. doi:10.1016/j.comppsych.2007.08.006

Diez-Quevedo, C., Rangil, T., Sanchez-Planell, L., Kroenke, K., & Spitzer, R. L. (2001).Validation and utility of

the Patient Health Questionnaire in diagnosing mental disorders in 1003 General Hospital Spanish Inpatients. *PsychosomaticMedicine, 63*, 679-686. doi:0033-3174/01/6304-0679

Donlan, W., & Lee, J. (2010).Screening for depression among indigenous Mexican migrant farmworkers using the Patient Health Questionnaire-9.*Psychological Reports, 106*, 419-432.

Dum, M., Pickren, J., Sobell, L. C., & Sobell, M. B. (2008). Comparing the BDI-II and the PHQ-9 with outpatient substance abusers. *Addictive Behaviors*, *33*(2), 381-387. doi:10.1016/j.addbeh.2007.09.017

Ell, K., Quon, B., Quinn, D. I., Dwight-Johnson, M., Wells, A., Lee, J., &Xie, B. (2007) Improving treatment for depression among low-income patients with cancer: The design of the ADAPt-C study. *General Hospital Psychiatry, 29*, 223-231. doi:10.1016/j.genhosppsych.2007.01.005

Fallowfield, L., Ratcliffe, D., Jenkins, V., & Saul, J. (2001).Psychiatric morbidity and its recognition by doctors in patients with cancer. *British Journal of Cancer, 84*, 1011-1015. doi:10.1054/bjoc.2001.1724

Fann, J. R., Berry, D. L., Wolpin, S., Austin-Seymour, M., Bush, N., Halpenny, B.,...McCorkle, R. (2009).Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer. *Psycho-oncology, 18*(1), 14-22. doi:10.1002/pon.1368

Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the patient health questionnaire (PHQ): A diagnostic metaanalysis. *Journal of General Internal Medicine, 22*(11), 1596-1602. doi:10.1007/s11606-007-0333-y

Granillo, M. T. (2012). Structure and function of the patient health questionnaire-9 among Latina and Non-Latina White female college students. *Journal of the Society for Social Work and Research, 3*(2), 80-93. doi:10.5243/jsswr.2012.6

Graves, D. E., & Bombardier, C. H. (2008). Improving the efficiency of screening for major depression in people with spinal cord injury. *The Journal of Spinal CordMedicine*, *31*(2), 177-184.

Gulec, M. Y., Gulec, H., Simsek, G., Turhan, M., & Sunbul, E. A. (2012).Psychometric properties of the Turkish version of the Patient Health Questionnaire–Somatic, Anxiety, and Depressive Symptoms. *Comprehensive Psychiatry, 53*(5), 623-629. doi:10.1016/j.comppsych.2011.08.002

Hansson, M., Chotai, J., Nordstom, A., & Bodlund, O. (2009). Comparison of two selfrating scales to detect depression: HADS and PHQ-9. *British Journal of General Practice*, *59*(566), e283-e288. doi:10.3399/bjgp09X454070

Hinz, A., Krauss, O., Hauss, J., Hockel, M., Kortmann, R., Stolzenburg, J., & Schwarz, R. (2010). Anxiety and depression in cancer patients compared with the general population. *European Journal of Cancer Care, 19*(4), 522-529. doi:10.1111/j.1365-2354.2009.01088.x.

Huang, F. Y., Chung, H., Kroenke, K., Delucchi, K. L., & Spitzer, R. L. (2006). Using the Patient Health Questionnaire‐9 to measure depression among racially and ethnically diverse primary care patients. *Journal of General InternalMedicine*, *21*(6), 547-552. doi: 10.1111/j.1525-1497.2006.00409.x

Johns, S. A., Kroenke, K., Krebs, E. E., Theobald, D. E., Wu, J., & Tu, W. (2013). Longitudinal comparison of three depression measures in adult cancer patients. *Journal of Pain and Symptom Management, 45*(1), 71-82. doi:10.1016/j.jpainsymman.2011.12.284

Kalpakjian, C. Z., Toussaint, L. L., Albright, K. J., Bombardier, C. H., Krause, J. K., & Tate, D. G. (2009). Patient Health Questionnaire-9 in spinal cord injury: an examination of factor structure as related to gender. *The Journal of Spinal Cord Medicine*, *32*(2), 147-156.

Karekla, M., Pilipenko, N., & Feldman, J. (2012). Patient Health Questionnaire: Greek language validation and subscale factor structure. *Comprehensive Psychiatry,53*(8), 1217-1226. doi:10.1016/j.comppsych.2012.05.008

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005).Prevalence, severity and comorbidity of 12-month DSM_IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 617-627. doi:10.1001/archpsyc.62.6.617

Krause, J. S., Reed, K. S., & McArdle, J. J. (2010). Factor structure and predictive validity of somatic and nonsomatic symptoms from the Patient Health Questionnaire-9: a longitudinal study after spinal cord injury. *Archives of physical medicine and rehabilitation*, *91*(8), 1218-1224. doi:10.1016/j.apmr.2010.04.015

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine, 16*, 606-613. doi:10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., Williams, J. B., & Lowe, B. (2010). The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. *General Hospital Psychiatry, 32*(4), 345-359. doi:10.1016/j.genhosppsych.2010.03.006

Lee, P. W., Schulberg, H. C., Raue, P. J., & Kroenke, K. (2007).Concordance between the PHQ-9 and the HSCL-20.*Journal of Affective Disorders, 99*, 139-145. doi:10.1016/j.jad.2006.09.002

Liu, S., Yeh, Z., Huang, H., Sun, F., Tjung, J., Hwang, L., …Yeh, A. W. (2011). Validation of Patient Health Questionnaire for depression screening among primary care patients in Taiwan.*Comprehensive Psychiatry, 52*(1), 96-101. doi:10.1016/j.comppsych.2010.04.013

Lotrakul, M., Sumrithe, S., & Saipanish, R. (2008).Reliability and validity of the Thai version of the PHQ-9.*BMC Psychiatry,8*, 46. doi:10.1186/1471-244X-8-46

Massie, M. J. (2004). Prevalence of depression in patients with cancer.*Journal of theNational Cancer Institute Monographs, 32*, 57-71. doi:10.1093/jncimonographs/lgh014

Ministerio da Saude (2004). *Plano Nacional de Saúde 2004/2010*. Lisboa: Ministerio da Saude.

Mitchell, A. J. (2007). Accuracy of Distress Thermometer and other ultra-short methods of detecting cancer-related mood disorders: Pooled results from 38 analyses. *Journal of Clinical Oncology, 25*, 1-12. doi:10.1200/JCO.2006.10.0438

Murray, C. J. L., & Lopez, A. D. (1996).*The global burden of disease: Summary*. Cambridge, MA: Harvard University.

Nunnally, J. C., & Bernstein, I. J. (1995). Teoria psicometrica. Madrid: McGraw-Hill.

Omoro, S. A. O., Fann, J. R., Weymuller, E. A., Macharia, I. M., & Yueh, B. (2006). Swahili translation and validation of the Patient Health Questionnaire-9 depression scale in the Kenyan head and neck cancer patient population. *The International Journal of Psychiatry in Medicine, 36*(3), 367-381.

Osorio, F. L., Mendes, A. V., Crippa, J. A., & Loureiro, S. R. (2009). Study of the discriminative validity of the PHQ-9 and PHQ-2 in a sample of Brazilian women in the context of primary health care. *Perspectives in Psychiatric Care, 45*(3), 216-227. doi:10.1111/j.1744-6163.2009.00224.x

Pais-Ribeiro, J., Pinto, C., & Santos, C. (2008). Validation study of the Portuguese version QLC-C30-V.3.*Psicologia, Saúde & Doenças, 9*(1), 89-102.

Pais-Ribeiro, J., Silva, I., Ferreira, T., Martins, A., Meneses, R., & Baltar, M. (2007). Validation study of a Portuguese version of the Hospital Anxiety and Depression Scale. *Psychology, Health & Medicine, 12*(2), 225-237. doi:10.1080/13548500500524088

Passik, S. D., Dugan, W., McDonald, M. V., Rosenfeld, B., Theobald, D. E., & Edgerton, S. (1998). Oncologists' recognition of depression in their patients with cancer.*Journal of Clinical Oncology, 16,* 1594-1600.

Pinto-Meza, A., Serrano-Blanco, A., Penarrubia, M. T., Blanco, E., & Haro, J. M. (2005). Assessing depression in primary care with the PHQ-9: can it be carried out over the telephone? *Journal of General Internal Medicine, 20*, 738-742. doi:10.1111/j.1525-1497.2005.0144.x

Pirl, W. F. (2010). Instruments in Psycho-oncology. In J. C. Holland, W. S. Breitbar, P. B. Jacobsen, M. S. Lederberg, M. J. Loscalzo, & R. McCorkle (Eds.), *Psychooncology* (pp. 119-130). New York: Oxford University Press.

Richardson, E. J., & Richards, J. S. (2008). Factor structure of the PHQ-9 screen for depression across time since injury among persons with spinal cordinjury. *Rehabilitation Psychology*, *53*(2), 243-249. doi:10.1037/0090-5550.53.2.243

Skarstein, J., Aass, N., Fossa, S. D., Skovlund, E., & Dahl, A. A. (2000). Anxiety and depression in cancer patients: relation between the Hospital Anxiety and Depression Scale and the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire. *Journal of Psychosomatic Research,49*(1), 27-34. doi:10.1016/S0022-3999(00)00080-5

Sollner, W., DeVries, A., Steixner, E., Lukas, P., Sprinzl, G., Rumpold, G., & Maislinger, S. (2001). How successful are oncologists in identifying patient distress, perceived social support, and need for psychosocial counselling? *British Journal Cancer, 84*, 179-185. doi:10.1054/bjoc.2000.1545

Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *JAMA, 282*(18), 1737-1744. doi:10.1001/jama.282.18.1737

Spitzer, R. L., Williams, J. B., Kroenke, K., Hornyak, R., & McMurray, J. (2000). Validity and utility of the Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetric-Gynecology Study. *American Journal of Obstetrics & Gynecology, 183*, 759-769. doi:10.1067/mob.2000.106580

Van't Spijker, A., Trjsburg, R. W., & Duivenvoorden, H. J. (1997). Psychological sequelae of cancer diagnosis: A meta-analytical review of 58 studies after 1980. *Psychosomatic Medicine, 59*, 280-293.

Vodermaier, A., Linden, W., & Siu, C. (2009). Screening for emotional distress in cancer patients: A systematic review of assessment instruments. *Journal of National Cancer Institute, 101*, 1464-1488. doi:10.1093/jnci/djp336

Wittkampf, K. A., Naeije, L., Schene, A. H., Huyser, J., & van Weert, H. C. (2007). Diagnostic accuracy of the mood module of the patient health questionnaire: A systematic review. *General Hospital Psychiatry, 29*, 388-395. doi:10.1016/j.genhosppsych.2007.06.004

Yeung, A., Fung, F., Yu, S., Vorono, S., Ly, M., Wu, S., & Fava, M. (2008).Validation of the Patient Health Questionnaire-9 for depression screening among Chinese Americans. Comprehensive Psychiatry, 49(2), 211-217. doi:10.1016/j.comppsych.2006.06.002

Yu, X., Tam, W. W., Wong, P. T., Lam, T. H., & Stewart, S. M. (2012). The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. Comprehensive Psychiatry, 53(1), 95-102. doi:10.1016/j.comppsych.2010.11.002

Zigmond, A. S., & Snaith, R. P. (1983).The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*, 361-370. doi:10.1111/j.1600-0447.1983.tb09716.x