Reliability and Score Ranges of the PHQ-9 and GAD-7 in a Primary and Secondary Care Mental Health Service

Article · December 2017

DOI: 10.1007/s40737-017-0090-0

CITATION
1

READS
781

4 authors, including:

Thomas Richardson
Solent NHS Trust
110 PUBLICATIONS **1,674** CITATIONS

SEE PROFILE

Mma Yeebo
Newcastle University
3 PUBLICATIONS **9** CITATIONS

SEE PROFILE

Journal of Psychosocial Rehabilitation and Mental Health

Reliability and Score Ranges of the PHQ-9 and GAD-7 in a Primary and Secondary Care Mental Health Service.

--Manuscript Draft--

Manuscript Number:	PRMH-D-17-00007R2
Full Title:	Reliability and Score Ranges of the PHQ-9 and GAD-7 in a Primary and Secondary Care Mental Health Service.
Article Type:	Brief report
Funding Information:	
Abstract:	Objectives: [C] The reliability of the PHQ-9 and GAD-7 as measures of depression and anxiety is well established in primary care. The present study evaluates whether or not differences occur for patients cared within a primary versus secondary mental health service. Methods: [C] Pre-treatment scores for patients were included in a cross-sectional cohort design. Results: [C] Reliability of the PHQ-9 and the GAD-7 ranged from "acceptable" to "excellent". Reliability and overall scores and did not significantly differ between either services. However many clients in both groups scored the maximum score on the GAD-7 and there were few differences in scores based on diagnosis for both measures. Conclusions: [C] The PHQ-9 and the GAD-7 are reliable measures of depression and anxiety symptoms for patients cared within a primary or secondary mental health service. However, the GAD-7 has a ceiling effect and both measures may not distinguish between services and diagnosis and therefore may not be useful for screening purposes.
Corresponding Author:	Thomas Richardson, D.Clin.Psych Solent NHS Trust Portsmouth, Hampshire UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Solent NHS Trust
Corresponding Author's Secondary Institution:	
First Author:	Thomas Richardson, D.Clin.Psych
First Author Secondary Information:	
Order of Authors:	Thomas Richardson, D.Clin.Psych
	Miles Wrightman
	Mma Yeebo
	Anna Lisicka
Order of Authors Secondary Information:	
Author Comments:	We have made the following changes: Regarding the test retest reliability: This study was cross-sectional so it is not possible to examine this. This has been added as a limitation in the discussion: "As the study was cross-sectional the test-retest reliability can also not be established."

Brief Report

Reliability and Score Ranges of the PHQ-9 and GAD-7 in a Primary and Secondary Care Mental Health Service

Thomas Richardson [a,b*], Miles Wrightman [c], MmaYeebo[a], Anna Lisicka[a]

[a] Mental Health Recovery Teams, Solent NHS Trust, St. Mary's Community Health Campus, Milton Road, Portsmouth, POE 6AD, U.K.

[b] School of Psychology, University of Southampton, Southampton, SO17 1BJ, U.K.

[c] Talking Change, 8F The Pompey Centre, Fratton Way, Portsmouth, PO4 8TA, U.K

Running Head: PHQ-9 and GAD-7 in Primary and Secondary Care

Keywords: Depression, Anxiety, Primary Care, Secondary Care.

* Corresponding Author: Dr Thomas Richardson, Email thr1g10@soton.ac.uk, Mental Health Recovery Team North, Solent NHS Trust, St. Mary's Community Health Campus, Milton Road, Portsmouth, POE 6AD, U.K. Phone (0044)7867461662.

Brief Report

Reliability and Score Ranges of the PHQ-9 and GAD-7 in a Primary and Secondary Care Mental Health Service [A]

Abstract [B]

Objectives: [C] The reliability of the PHQ-9 and GAD-7 as measures of depression and anxiety is well established in primary care. The present study evaluates whether or not differences occur for patients cared within a primary versus secondary mental health service.

Methods: [C] Pre-treatment scores for patients were included in a cross-sectional cohort design.

Results: [C] Reliability of the PHQ-9 and the GAD-7 ranged from "acceptable" to "excellent". Reliability and overall scores did not significantly differ between services. However many clients in both groups scored the maximum score on the GAD-7 and there were few differences in scores based on diagnosis for both measures.

Conclusions: [C] The PHQ-9 and the GAD-7 are reliable measures of depression and anxiety symptoms for patients cared within a primary or secondary mental health service. However, the GAD-7 has a ceiling effect and both measures may not distinguish between services and diagnosis and therefore may not be useful for screening purposes.

Keywords: [C] Depression; Anxiety; Primary Care; Secondary Care.

Introduction [B]

The Patient Health Questionnaire (PHQ-9; (Kroenke, Spitzer, & Williams, 2001) and Generalised Anxiety Disorder Questionnaire (GAD-7; (Spitzer, Kroenke, Williams, & Löwe, 2006) are commonly used measures of depression and anxiety symptoms in primary and secondary care mental health services in the United Kingdom (UK). Both measures are also frequently administered to patients diagnosed with various physical health conditions. These are used both for initial screening and as an outcome measure.

The validity and reliability of the PHQ-9 and GAD-7 was first established by Kroenke et al. (2001) and Spitzer et al. (2006) principally within primary care mental health services. The PHQ-9 demonstrated 88% sensitivity and specificity, whilst the GAD-7 revealed 89% sensitivity and 82% specificity. The PHQ-9 and GAD-7 showed good reliability (Cronbach's alpha =.89; .92, respectively). Both measures were suggested to be effective in screening for and monitoring symptoms depression and generalized anxiety.

However, recent studies in secondary care have suggested good reliability but poor specificity. Kertz, Bigda-Peyton, and Bjorgvinsson (2013) examined the validity of the GAD-7 in inpatients finding good internal consistency and sensitivity suggesting it is a useful measure of anxiety symptoms in this sample, however that it had poor specificity in identifying patients with GAD. Beard and Björgvinsson (2014) similarly found good internal consistency of the GAD-7 but poor specificity and suggested it as a measure of anxiety but not screen for anxiety disorders in a heterogeneous sample.

To our knowledge, no studies have compared the reliability and scores ranges of the PHQ-9 and GAD-7 between patients cared within a primary versus secondary mental health service.

Method [B]

Design & Procedure [C]

Pre-treatment scores for patients cared within a primary versus secondary mental health service were included in a cross-sectional cohort design. Both services were National Health Services in Portsmouth, UK. Measures were collected during routine clinical practice (i.e., at assessment or first therapy session), thus the local Trust Research Department approved the study as a service evaluation. This means that as these questionnaires were completed for routine clinical practice, ethics approval as a full research project was not needed as clients were not being asked to do anything outside of their usual care. Approval was therefore to analyze this pre-existing data for a service evaluation. **Participants [C]**

The dataset consisted of 387 patients; 258 patients (56.2% female; *M* age = 37.9 years, age range: 18–66) from primary care and 129 patients from secondary care (71.3% female; *M* age= 37.1 years, age range: 18–65). Patients were majority 'White European' ethnicity in primary care (97.1%) and secondary care (93.2%). Table 1 presents the primary diagnoses of patients from primary versus secondary care. The most frequent diagnosis was major depressive disorder in primary care (29.5%) and emotionally unstable personality disorder in secondary care (30.2%). A secondary diagnosis was more prevalent in secondary care (48%) versus primary care (28%).

Measures [C]

PHQ-9. [D] A self-report scale consisting of nine items that measure the frequency of depression symptoms (e.g., feeling hopeless, little interest or pleasure in doing things, negative self-evaluation). Scores range from 0–27, and are classified 'mild' (5–9), 'moderate' (10–14), 'moderately severe' (15–19) and 'severe' (20–27). The clinical range is declared ≥ 10.

GAD-7. [D] A self-report scale consisting of seven items that measure the frequency of generalized anxiety symptoms (e.g., feeling nervous, being unable to stop or control worrying, becoming easily annoyed or irritable). Scores range from 0–21, and are classified

'mild' (5–9), 'moderate' (10–14) and 'severe' (15–21). The clinical range is declared ≥ 8. Respondents are asked to retrospectively assess the frequency of their symptoms over the previous seven (primary care) or fourteen days (secondary care).

Results [B]

Completers & Missing items [C]

The PHQ-9 and GAD-7 was completed by 92.2% and 99.6% of patients cared within a primary mental health service versus 55% and 65.9% of patients cared within a secondary mental health service, respectively[1]. Missing items (.5–1.3%) were transformed into the sample mean.

Preliminary analysis [C]

Patients cared within primary (M = 19.3, SD = 6.2) versus secondary care (M = 17.7, SD = 5.6) did not differ significantly on PHQ-9 scores, $t(306) = 1.932$, $p > .05$ Additionally, patients cared within primary (M = 15.6, SD = 5.4) versus secondary care (M = 15.9, SD = 4.6) did not differ significantly on GAD-7 scores, $t(340) = -.540$, $p = .589$. Maximum scores on the PHQ-9 and GAD-7 were attained by 9.3% and 17.1% of primary care patients versus 2.8% and 15.3% of secondary care patients, respectively.

Reliability [C]

Chronbach's Alpha Reliability of the PHQ-9 was classified "acceptable" (α = .79) and "good" (α = .81) for primary versus secondary care patients, respectively. Moreover, Chronbach's Alpha Reliability of the GAD-7 was classified "excellent" (α = .9) and "good" (α = .84) for primary versus secondary care patients, respectively.

Differences Based on Diagnosis

Primary care patients with an anxiety diagnosis (M = 17.6, SD = 6.2) scored significantly higher on the GAD-7 than patients with another or no anxiety diagnosis (M =

[1] The PHQ-9 and GAD-7 were routinely administered together in primary care whereas in secondary care either or both measures were administered dependant on the type of psychotherapy to which the patient was assigned.

15.2, *SD* = 5.7), *t*(255) = 2.62, *p* = <.01. No significant differences on the GAD-7 were identified for secondary care patients with an anxiety disorder diagnosis (*M* = 13.7, *SD* = 2.9) and patients without identified diagnosis (*M* = 16.24, *SD* = 4.9), *t*(83) = -1.61, *p* = >.05 . Additionally, no significant differences on the PHQ-9 were identified for primary care patients with major depressive disorder or bipolar disorder (*M* = 19.5, *SD* = 5.6) and patients with no identified diagnosis (*M* = 19.1, *SD* = 6.5), *t*(235) = 0.49, *p* = >.05. Moreover, no significant differences on the PHQ-9 / were identified for secondary care patients with major depressive disorder or bipolar disorder (*M* = 16.9, *SD* = 6.2) and without either diagnosis (*M* = 19.1, *SD* = 6.5), *t*(69) = -1.58, *p* = >.05.

Discussion[B]

The study provides evidence for reliability of the PHQ-9 and GAD-7 as measures of anxiety and depression symptoms. This applied to both primary and secondary mental health services. The data suggested that both measures were effective in assessing symptom severity in primary and secondary mental health in line with the work of Kertz et al. (2013) and Beard and Björgvinsson (2014) who suggest the GAD-7 is reliable for measures anxiety severity in secondary care. This paper adds to this that the PHQ-9 is a useful measure of depression severity in both settings. There was little missing data which shows good acceptability of the measures, perhaps due to them being short. However, it is important to note that though clients are not required to complete these questionnaires in order to receive a service, they may have felt pressured to do so. Both settings had overall mean scores greater than the recommended cut-off points for identifying cases of anxiety and depression. However no difference was found in scores between secondary and primary mental health, which suggests that the measures may not representative of overall symptom severity or complexity, and should not be used as a screen to distinguish whether primary or secondary care services are most appropriate for psychological therapy. It is perhaps surprising that mean scores for the

measures were in the severe range. However high intensity Increasing Access to Psychological therapies (IAPT) services in the UK are meant to see moderate to severe depression (Clark, 2011). In addition, a study from the first year of IAPT found that those with mixed depression and anxiety had pre mean PHQ-9 scores of 16 and mean GAD-7 scores of 14 (Gyani, Shafran, Layard & Clark, 2013), compared to 19 and 15 in the current study. Therefore scores may be only slightly higher than the national picture. This may be because Portsmouth has higher deprivation and poorer physical health than the average in England (Public Health England, 2014). It may also be that this specific service has increased its intake of more severe/complex cases in recent years.

Though no formal specificity analysis was carried out, scores distinguished anxiety disorders in primary care only. Therefore GAD-7 did not present as an efficient screening tool for anxiety disorders in secondary services, in line with previous findings (Beard & Björgvinsson, 2014; Kertz et al., 2013). The PHQ-9 did not distinguish depression from other difficulties in either primary or secondary care, suggesting it is useful as a measure of symptoms severity and outcome measure in both settings, but may not be reliable for screening or diagnostic purposes.

A possible ceiling effect was identified for GAD-7 particularly in both settings. This may suggest limited sensitivity of the tool. However it is well recognized that screening tools for anxiety disorders are usually less sensitive than screening tools for depressive disorders (Rose & Devine, 2014). This may be explained by the fact that different types of anxiety disorders present with greater heterogeneous symptoms than different types of depressive disorders.

Several limitations from this study should be noted. As previously discussed, the study did not examine specificity formally. As the study was cross-sectional the test-retest reliability can also not be established. A major limitation is that this was a service evaluation

of existing data rather than a research project, which might have evaluated validity in a more rigorous way for example comparing questionnaires to scores on structured clinical interviews. There was limited ethnic diversity, and lower numbers of secondary care participants.

In conclusion, the PHQ-9 and GAD-7 are useful measures of symptom severity in both primary and secondary care, but may not be effective as screening tools.

Acknowledgements [C]

Thank you to the voluntary research assistants who helped with data entry, and Sue Sadler for helping with collating questionnaires.

Conflict of interest [C]

The authors have no conflicts of interest to declare.

References [B]

Beard C, Björgvinsson T. Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. Journal of Anxiety Disorders. 2014; 28(6): 547-552.

Clark, DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. International Review of Psychiatry. 2011; 23: 318–327.

Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: Lessons from year one of IAPT. Behaviour Research and Therapy. 2013; 51, 597-606.

Kertz S, Bigda-Peyton J, Bjorgvinsson T. Validity of the Generalized Anxiety Disorder-7 Scale in an acute psychiatric sample. Clinical Psychology & Psychotherapy. 2013; 20(5): 456-464.

Kroenke K, Spitzer RL, Williams JB. The PHQ-9. Journal of General Internal Medicine. 2001; 16(9): 606-613.

NHS England. Portsmouth Health Profile. 2014. Obtained from Portsmouth Clinical Commissioning Group Website (www.portsmouthccg.nhs.uk*)* on 02/05/2017.

Rose M, Devine J. Assessment of patient-reported symptoms of anxiety. Dialogues in Clinical Neuroscience. 2014; 16(2): 197-211.

Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine. 2006; 166(10); 1092-1097.

Table 1.

Primary Diagnoses of Patients Cared by a Primary versus Secondary Mental Health Service

Primary Diagnosis	Primary Care		Secondary Care	
	n	*(%)*	*n*	*(%)*
No Diagnosis	81	(31.4)	0	(0)
Depression	76	(29.5)	34	(26.4)
Bipolar Disorder	12	(4.7)	12	(9.3)
Post-Traumatic Stress Disorder	8	(3.1)	1	(.8)
Emotionally Unstable Personality Disorder	11	(4.3)	39	(30.2)
Other Personality Disorder	21	(8.2)	3	(2.4)
Obsessive-Compulsive Disorder	6	(2.3)	7	(5.4)
Generalised Anxiety Disorder	18	(7.3)	5	(3.9)
Panic and/or Agoraphobia	4	(1.6)	1	(.8)
Psychosis	10	(3.9)	4	(3.1)
Alcohol Abuse or Dependence (current)	5	(1.9)	3	(2.3)
Phobia	6	(2.3)	1	(.8)
Other Diagnosis	0	(0)	18	(13.9)

We have made the following changes:

Regarding the test retest reliability: This study was cross-sectional so it is not possible to examine this. This has been added as a limitation in the discussion: “As the study was cross-sectional the test-retest reliability can also not be established.”