Journal of Affective Disorders 232 (2018) 34–40

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

The PHQ-9 Item 9 based screening for suicide risk: a validation study of the Patient Health Questionnaire (PHQ)−9 Item 9 with the Columbia Suicide Severity Rating Scale (C-SSRS)

Peter J. Na[a,*], Satyanarayana R. Yaramala[a], Jihoon A. Kim[b], Hyelee Kim[b], Fernando S. Goes[c], Peter P. Zandi[c], Jennifer L. Vande Voort[a], Bruce Sutor[a], Paul Croarkin[a], William V. Bobo[a]

[a] Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA
[b] Seoul National University College of Medicine, Seoul, Republic of Korea
[c] Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

ARTICLE INFO

Keywords:
PHQ-9
C-SSRS
Suicidal ideation
Suicide risk
Validation

ABSTRACT

Background: Item 9 of the Patient Health Questionnaire (PHQ) evaluates passive thoughts of death or self-injury within the last two weeks, and is often used to screen depressed patients for suicide risk. We aimed to validate the PHQ-9 item 9 with a brief electronic version of the Columbia Suicide Severity Rating Scale (eC-SSRS).

Methods: We analyzed data from 841 patients enrolled in the National Network of Depression Centers Clinical Care Registry. We performed a validation analysis of PHQ-9 item 9 for suicide risk and ideation, using the eC-SSRS as a gold standard (defined as positive response to suicidal ideation with intent to act or recent suicidal behavior).

Results: Of the 841 patients, 13.4% and 41.1% were assessed as being positive for suicide risk by the eC-SSRS and PHQ-9 item 9, respectively. For the overall cohort, sensitivity was 87.6% (95%CI 80.2–92.5%), specificity was 66.1% (95%CI 62.6–69.4%), PPV was 28.6% (95%CI 24.1–33.6%), and NPV was 97.2% (95%CI 95.3–98.3%) for the PHQ-9 suicide item. These performance measures varied within subgroups defined by demographic and clinical characteristics. In addition, the validity of PHQ-9 item 9 (cutoff score of 1) with eC-SSRS-defined suicide ideation showed overall poor results.

Limitations: The gold standard used in our study was a surrogate measure of suicidality based on eC-SSRS scores.

Conclusions: The results of our study suggest that item 9 of the PHQ-9 is an insufficient assessment tool for suicide risk and suicide ideation, with limited utility in certain demographic and clinical subgroups that requires further investigation.

1. Introduction

Suicide rates in the U.S. have increased during the past 15 years, peaking at 13.0 cases per 100,000 population in 2014 as compared to 10.5 cases per 100,000 population in 1999 (Center for Disease Control and Prevention, 2016). Suicide has gained increasing attention as a major public health concern. Notably, suicide completers have a significantly higher number of outpatient, inpatient and emergency department mental health visits prior to death as compared to age- and sex-matched controls (Chock et al., 2015). As such, screening depressed patients for suicidal ideation and intent has been recommended for routine in practice in primary care physicians' offices, emergency departments, and psychiatry outpatient clinics (Conway et al., 2016; Joint Commissions, 2016; Luoma et al., 2002; Uebelacker et al., 2011; Valuck et al., 2012).

The 9-item Patient Health Questionnaire (PHQ-9) has been shown to be an effective screening tool for major depression and a reliable and valid measure of depressive symptom severity in a variety of clinical settings (Gilbody et al., 2007; Valuck et al., 2012; Wittkampf et al., 2007). Item 9 of the PHQ-9 (referred to hereafter as the 'PHQ-9 suicide item') specifically evaluates the frequency of passive thoughts of death or self-injury within the last two weeks, and has been widely used as a single measure to assess the prevalence of suicidal ideation in research studies (Bauer et al., 2013; Denneson et al., 2014; Walker et al., 2011, 2008; Yawn et al., 2009). In addition, the PHQ-9 suicide item has been used as a brief screening measure for suicide risk (Simon et al., 2013). However, this practice is controversial because the predictive value of the PHQ-9 suicide item for suicide risk is unclear (Bauer et al., 2013;

* Correspondence to: 550 First Ave, New York, NY 10016, USA.
E-mail address: peter.na@nyumc.org (P.J. Na).

https://doi.org/10.1016/j.jad.2018.02.045
Received 12 May 2017; Received in revised form 14 February 2018; Accepted 16 February 2018
Available online 17 February 2018
0165-0327/

Table 1
Demographic and clinical characteristics based on eC-SSRS and PHQ-9 Item 9 responses.

Variables	eC-SSRS (+)[a]		eC-SSRS (-)[b]			PHQ Item 9 (+)		PHQ Item 9 (-)		
	n	%	*n*	%	*p-value*	*n*	%	*n*	%	*p-value*
Total	113	13.4	728	86.6		346	41.1	495	58.9	
Female sex	69	61.1	481	66.1	0.30	210	60.7	340	68.7	0.02
Age strata, years					0.01					0.002
19–49	77	68.1	398	54.7		206	59.5	269	54.3	
50–64	25	22.1	187	25.7		96	27.7	116	23.4	
65 or above	11	9.7	143	19.6		44	12.7	110	22.2	
Race					0.064					0.215
White	101	89.4	600	82.4		295	85.3	406	82.0	
Non-white	12	10.6	128	17.6		51	14.7	89	18.0	
Marital status					0.04					0.46
Single (Never married)	31	27.7	218	30.2		116	33.7	133	27.1	
Married or in a committed relationship	45	40.2	349	48.3		145	42.2	249	50.8	
Separated, widowed or divorced	36	32.1	155	21.5		83	24.1	108	22.0	
Living alone	35	31.3	202	27.9		95	27.5	142	28.9	0.66
Education					0.008					< 0.001
High school degree or less	24	21.4	119	16.5		75	21.9	68	13.9	
Technical/Associates degree or some college	50	44.6	234	32.5		139	40.5	145	29.7	
Bachelor's degree	18	16.1	200	27.8		78	22.7	140	28.6	
Advanced/Professional degree	20	17.9	167	23.2		51	14.9	136	27.8	
At least one psychiatric hospitalization	92	81.4	361	50.9	< 0.001	227	66.1	226	47.2	< 0.001
Pregnant women	2	2.9	44	9.1	< 0.001	10	4.8	36	10.6	< 0.001
Primary mood disorder diagnosis					0.001					0.03
Bipolar disorder, any type	19	17.3	230	34.6		91	27.7	158	35.4	
Major depressive disorder	68	61.8	353	53.1		184	55.9	237	53.1	
Other mood disorder	23	20.9	82	12.3		54	16.4	51	11.4	
Treatment resistance[c]	61	54.5	351	50.9	0.49	163	48.3	249	53.7	0.14
PHQ−8 score categories at initial visit					< 0.001					< 0.001
No depression (score of 9 or below)	20	17.7	308	42.3		50	14.5	278	56.2	
Mild depression (score of 10–19)	56	49.6	299	41.1		171	49.4	184	37.2	
Major depression (score of 20 or above)	37	32.7	121	16.6		125	36.1	33	6.7	
	Mean	**SD**	**Mean**	**SD**	**p-value**	**Mean**	**SD**	**Mean**	**SD**	**p-value**
Age, years	42.1	12.9	46.1	16.3	0.02	44.4	15.2	46.4	16.4	0.04
Lifetime psychiatric hospitalizations (n)	2.2	1.8	1.3	1.7	< 0.001	1.8	1.8	1.1	1.6	< 0.001
PHQ−8 total score	15.7	5.9	10.7	6.7	< 0.001	15.7	5.4	8.5	6.0	< 0.001

[a] Defined as having suicidal ideation with intent to act or recent suicidal behavior, i.e., any affirmative response to eC-SSRS items 5 or 6b or 7b.
[b] Defined as absence of suicidal ideation with intent to act or recent suicidal behavior, i.e., no affirmative response to any of eC-SSRS items 5 or 6b or 7b.
[c] Treatment resistance is defined as participants who have failed to respond or had a relapse of depressive symptoms despite receiving at least one adequate course of antidepressant treatment.

Louzon et al., 2016; Razykov et al., 2012a, 2012b; Simon et al., 2013; Uebelacker et al., 2011; Viguera et al., 2015; Walker et al., 2011, 2010). Moreover, the PHQ-9 suicide item assesses both passive thoughts of death and the desire for self-harm in a single response item, which may yield high false-positive rates (Walker et al., 2011) –an important clinical consideration given the propensity for false-positive suicide risk screenings to cause undue stress upon the limited existing mental health resources (Thompson et al., 2004), which could result in unnecessary urgent assessments and emergency department visits (Weil, 2015).

The Columbia-Suicide Severity Rating Scale (C-SSRS) is a reliable and valid scale that assesses a broader spectrum and severity of suicidal ideation and behavior, including suicidal intent (Conway et al., 2016; Gill et al., 2015; Kerr et al., 2014; Madan et al., 2016; Posner et al., 2011; United States Food and Drug Administration, 2012; Youngstrom et al., 2015). Some have proposed that the C-SSRS be used to screen for suicidal ideation and intent prior to conducting a more in-depth evaluation of suicide risk (Scott et al., 2009; Walker et al., 2011). Recently, Viguera and colleagues compared the utility of PHQ-9 suicide item, an abbreviated self-report version of the full-scale C-SSRS, and clinical assessment of suicide risk in a cohort of 1620 adults in a tertiary care, psychiatric outpatient clinic and reported high false positive rates generated by PHQ-9 suicide item compared to a C-SSRS based definition of elevated suicide risk (any affirmative response to item 5) self-reported suicidal ideation with intent to act in the past month, or 6b) actual, aborted or interrupted suicide attempts within the past 3 months, or 7b) a behavior preparatory for making an attempt within the past 3 months were classified as positive for suicide risk). Additional investigations of the validity of the PHQ-9 suicide item are needed in mental health specialty settings, where patients with prior suicide attempts and active suicidal ideation are most likely to be found (Gaynes et al., 2007; Schmitz et al., 2012). The aim of our study was to carry out a similar validity analysis of the PHQ-9 suicide item using the C-SSRS as a gold standard (using the aforementioned criteria of Viguera et al. (2015)), and extend this work by assessing the validity of the PHQ-9 suicide item with the C-SSRS suicide ideation item.

We thus conducted a cross-sectional analysis of data from the National Network of Depression Centers (NNDC) Clinical Care Registry (CCR) to measure the positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity of the PHQ-9 suicide item, using definitions of suicide ideation with intent and suicide ideation derived from a brief, electronic version of the full-scale C-SSRS as gold standards. We hypothesized that the PPVs of the PHQ-9 suicide item would be poor in a cohort of patients evaluated in academic psychiatric clinics specializing in the treatment of depression.

2. Method

2.1. Cohort

The NNDC-CCR is a long term, prospective, observational, multicenter registry with an affiliation of 15 U.S.-based academic medical centers specializing in the assessment and treatment of depression. Details of the registry's sampling methods have been published previously (Forester et al., 2014; Weiss et al., 2016).