INSTRUCTION MANUAL
Instructions for Patient Health Questionnaire (PHQ) and GAD-7 Measures

TOPIC	PAGES
Background	1
Coding and Scoring	2, 4, 5
Versions	3
Use as Severity and Outcome Measures	6-7
Translations	7
Website and Other Issues	8
Selected References	9

BACKGROUND

The Primary Care Evaluation of Mental Disorders (PRIME-MD) was an instrument developed and validated in the early 1990s to efficiently diagnose five of the most common types of mental disorders presenting in medical populations: depressive, anxiety, somatoform, alcohol, and eating disorders.[1] Patients first completed a one-page 27-item screener and, for those disorders for which they screened positive, were asked additional questions by the clinician using a structured interview guide. However, this 2-stage process took an average of 5-6 minutes of clinician time in patients without a mental disorder diagnosis and 11-12 minutes in patients with a diagnosis. This proved to be a barrier to use given the competing demands in busy clinical practice settings.

Therefore, in two large studies enrolling 6000 patients (3000 from general internal medicine and family practice clinics and 3000 from obstetrics-gynecology clinics), a self-administered version of the PRIME-MD called the Patient Health Questionnaire (PHQ) was developed and validated.[2,3] In the past decade, the PHQ in general and the PHQ-9 depression scale in particular [4-6] have gained increasing use in both research and practice. The original PRIME-MD is now largely of historical interest and seldom used except in a few types of research studies.

Given the popularity of the PHQ-9 for assessing and monitoring depression severity, a new 7-item anxiety scale using a response set similar to the PHQ-9 was initially developed to diagnose generalized anxiety disorder (hence its name, the GAD-7) and validated in 2740 primary care patients.[7] Though originally developed to diagnose generalized anxiety disorder, the GAD-7 also proved to have good sensitivity and specificity as a screener for panic, social anxiety, and post-traumatic stress disorder.[8] Finally, the PHQ-15 was derived from the original PHQ studies and is increasingly used to assess somatic symptom severity and the potential presence of somatization and somatoform disorders.[9]

Each PHQ module can be used alone (e.g. the PHQ-9 if depression is the condition of interest), together with other modules, or as part of the full PHQ. Also, alternative or abbreviated versions of the PHQ-9 and GAD-7 are sometimes used in certain screening or research settings [10-14]
Although the PHQ was originally developed to detect five disorders, the depression, anxiety, and somatoform modules (in that order) have turned out to be the most popular.[10] Also, most primary care patients with depressive or anxiety disorders present with somatic complaints and co-occurrence of somatic, anxiety, and depressive symptoms (the *SAD* triad) is exceptionally common. This is the rationale behind the PHQ-SADS screener.[15] The most commonly used versions of the PHQ scales are summarized in **Table 1, page 3**.

CODING AND SCORING

The full PHQ, Brief PHQ, and PHQ for Adolescents (PHQ-A) can be used to establish provisional diagnoses for selected DSM-IV disorders. The diagnostic algorithm for the PHQ modules are included in footers at the bottom of each page of the PHQ, and also reiterated in **Table 2, page 4**. The other measures are principally used to derive severity scores (PHQ-9 and PHQ-8 for depressive symptom severity; GAD-7 for anxiety symptom severity; PHQ-15 for somatic symptom severity) or as ultra-brief screeners (PHQ-2, GAD-2, PHQ-4). An example in which the PHQ depression module can be used as both a diagnostic module as well as a depression severity score (PHQ-9 score) is shown in **Table 3, page 5.**

Over time, the severity scores have been a particularly popular use of the measures, and are now used much more commonly than the provisional diagnoses. For example, cutpoints of 5, 10, and 15 represent mild, moderate, and severe levels of depressive, anxiety, and somatic symptoms, on the PHQ-9, GAD-7, and PHQ-15 respectively. Also, a cutpoint of 10 or greater is considered a “yellow flag” on all 3 measures (i.e., drawing attention to a possible clinically significant condition), while a cutpoint of 15 is a “red flag” on all 3 measures (i.e., targeting individuals in whom active treatment is probably warranted). For the ultra-brief measures (PHQ-2 and GAD-2), a score of 3 or greater should prompt administration of the full PHQ-9 and/or GAD-7, as well as a clinical interview to determine whether a mental disorder is present.

The final question on the PHQ (and some of its abbreviated versions) asks the patients to report “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?” This single patient-rated difficulty item is not used in calculating any PHQ score or diagnosis but rather represents the patient’s global impression of symptom-related impairment. It may be useful in decisions regarding initiation of or adjustments to treatment since it is strongly associated with both psychiatric symptom severity as well as multiple measures of impairment and health-related quality of life.

A particularly important question is how to assess suicide risk in individuals who answer positively to the 9th question of the PHQ-9. A four-item screener has been developed that may assist in positive responses to this 9th question [16], although a final decision about the actual risk of self-harm requires a clinical interview.

Table 1. Versions: Patient Health Questionnaire (PHQ) Family of Measures

Measure	Description	Scoring	References
Core			
PRIME-MD	Predecessor of PHQ, now mainly of historical interest.	Combined self-administered patient screener with clinician follow-up questions.	1
PHQ	Five modules covering 5 common types of mental disorders: depression, anxiety, somatoform, alcohol, and eating.	Selected (but provisional) DSM-IV diagnoses for all types of disorders except somatoform.	2, 3
PHQ-9	Depression scale from PHQ.	Nine items, each of which is scored 0 to 3, providing a 0 to 27 severity score.	1, 4, 5, 6, 10
GAD-7	Anxiety measure developed after PHQ but incorporated into PHQ-SADS.	Seven items, each of which is scored 0 to 3, providing a 0 to 21 severity score.	7, 8, 10
PHQ-15	Somatic symptom scale from PHQ.	Fifteen items, each of which is scored 0 to 2, providing a 0 to 30 severity score.	9, 10
PHQ-SADS	PHQ-9, GAD-7, and PHQ-15 measures, plus panic measure from original PHQ.	See scoring for these scales above.	10
Variants			
Brief PHQ	PHQ-9 and panic measures from original PHQ plus items on stressors and women's health.	See scoring for PHQ above. Stressor and women's health items are not diagnostic or scored.	3
PHQ-A	Substantially modified version of PHQ developed for use in adolescents. Moderate data exists for validity but much less than for original PHQ.	Diagnostic scoring described in manual, available upon request.	11
PHQ-2	First 2 items of PHQ-9. Ultra-brief depression screener.	Two items scored 0 to 3 (total score of 0-6)	10, 12
GAD-2	First 2 items of GAD-7. Ultra-brief anxiety screener.	Two items scored 0 to 3 (total score of 0-6)	8, 10, 12
PHQ-4	PHQ-2 and GAD-2.	See PHQ-2 and GAD-2 above.	10, 12, 13
PHQ-8	All items of PHQ-9 except the 9th item on self-harm. Mainly used in non-depression research studies.	Eight items, each of which is scored 0 to 3, providing a 0 to 24 severity score.	5, 10, 14

Table 2. Diagnostic Algorithms for the PHQ

Page 1

Somatoform Disorder if at least 3 of #1a-m bother the patient "a lot" and lack an adequate biological explanation.

Major Depressive Syndrome if #2a or b and five or more of #2a-i are at least "More than half the days" (count #2i if present at all) .

Other Depressive Syndrome if #2a or b and two, three, or four of #2a-i are at least "More than half the days" (count #2i if present at all).

Note**:** the diagnoses of Major Depressive Disorder and Other Depressive Disorder requires ruling out normal *bereavement (mild symptoms, duration less than 2 months***)**, a history of a *manic* episode (Bipolar Disorder) and a *physical disorder, medication or other drug* as the biological cause of the depressive symptoms.

Page 2

Panic Syndrome if #3a-d are all 'YES' and 4 or more of #4a-k are 'YES'.

Other Anxiety Syndrome if #5a and answers to three or more of #5b-g are "More than half the days".

Note**:** The diagnoses of Panic Disorder and Other Anxiety Disorder require ruling out a *physical disorder***,** *medication or other drug* as the biological cause of the anxiety symptoms.

Page 3

Bulimia Nervosa if #6a,b, and c and #8 are 'YES';

Binge Eating Disorder the same but #8 is either 'NO' or left blank.

Alcohol abuse if any of #10a-e are "YES".

Additional Clinical Considerations. After making a provisional diagnosis with the PHQ, there are additional clinical considerations that may affect decisions about management and treatment.

- *Have current symptoms been triggered by psychosocial **stressor(s**)?*
- *What is the **duration** of the current disturbance and has the patient received any **treatment** for it?*
- *To what extent are the patient's symptoms i**mpairing** his or her usual work and activities?*
- *Is there a **history** of similar episodes, and were they **treated**?*
- *Is there a **family history** of similar conditions?*

Table 3. Example of PHQ Depression Module for both Diagnostic and Severity Purposes

Patient: A 43-year-old woman who looks sad and complains of fatigue for the past month.

2. Over the last 2 weeks, how often have you been bothered by any of the following:	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
	(0)	(1)	(2)	(3)
a. Little interest or pleasure in doing things?……..	☐	☐	☐	☒
b. Feeling down, depressed, or hopeless?……..…	☐	☒	☐	☐
c. Trouble falling or staying asleep, or sleeping too much?	☐	☐	☒	☐
d. Feeling tired or having little energy?.....……...…..	☐	☐	☐	☒
e. Poor appetite or overeating?.....…….….	☐	☒	☐	☐
f. Feeling bad about yourself—or that you are a failure or have let yourself or your family down?……..	☐	☐	☒	☐
g. Trouble concentrating on things, such as reading the newspaper or watching television? ……..	☐	☐	☐	☒
h. Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual?.....……..	☒	☐	☐	☐
i. Thoughts that you would be better off dead or of hurting yourself in some way? …….	☐	☒	☐	☐

FOR OFFICE CODING: Maj Dep Syn if #2a or b and five or more of #2a-i are at least "More than half the days" (count #2i if present at all) . Other Dep Syn if #2a or b and two, three, or four of #2a-i are at least "More than half the days" (count #2i if present at all).

Major Depressive Disorder Diagnosis. The criteria for Major Depressive Syndrome are met since she checked #2a "nearly every day" and five of items #2a to i were checked "more than half the days" or "nearly every day". Note that #2i, suicidal ideation, is counted whenever it is present.

In this case, the diagnosis of Major Depressive Disorder (not Syndrome) was made since questioning by the physician indicated no history of a manic episode; no evidence that a physical disorder, medication, or other drug caused the depression; and no indication that the depressive symptoms were normal bereavement. Questioning about the suicidal ideation indicated no significant suicidal potential.

PHQ-9 Depression Severity. This is calculated by assigning scores of 0, 1, 2, and 3, to the response categories of "not at all," "several days," "more than half the days," and "nearly every day," respectively. PHQ-9 total score for the nine items ranges from 0 to 27. In the above case, the PHQ-9 depression severity score is 16 (3 items scored 1, 2 items scored 2, and 3 items scored 3). Scores of 5, 10, 15, and 20 represent cutpoints for mild, moderate, moderately severe and severe depression, respectively. Sensitivity to change has also been confirmed.

USE OF SOME SCREENERS AS SEVERITY AND OUTCOME MEASURES

PHQ-9 Depression Severity. This is calculated by assigning scores of 0, 1, 2, and 3, to the response categories of "not at all," "several days," "more than half the days," and "nearly every day," respectively. PHQ-9 total score for the nine items ranges from 0 to 27. In the above case (see table 3, page 5), the PHQ-9 depression severity score is 16 (3 items scored 1, 2 items scored 2, and 3 items scored 3). Scores of 5, 10, 15, and 20 represent cutpoints for mild, moderate, moderately severe and severe depression, respectively. Sensitivity to change has also been confirmed. The **PHQ-8** is scored just like the PHQ-9 and its total score ranges from 0 to 24. Cutpoints on the PHQ-8 are identical to the PHQ-9.

GAD-7 Anxiety Severity. This is calculated by assigning scores of 0, 1, 2, and 3, to the response categories of "not at all," "several days," "more than half the days," and "nearly every day," respectively. GAD-7 total score for the seven items ranges from 0 to 21. Scores of 5, 10, and 15 represent cutpoints for mild, moderate, and severe anxiety, respectively. Though designed primarily as a screening and severity measure for generalized anxiety disorder, the GAD-7 also has moderately good operating characteristics for three other common anxiety disorders – panic disorder, social anxiety disorder, and post-traumatic stress disorder. When screening for anxiety disorders, a recommended cutpoint for further evaluation is a score of 10 or greater.

PHQ-2 and GAD-2 Severity. These consist of the first two items of the PHQ-9 and GAD-7 respectively, and constitute the two core DSM-IV items for major depressive disorder and generalized anxiety disorder, respectively. Each ranges from a score of 0 to 6. The operating characteristics of these ultra-brief measures are quite good; the recommended cutpoints for each when used as screeners is a score of 3 or greater. When used together, they are referred to as the **PHQ-4** a 4-item screening measure which ranges from a score of 0 to 12, and serves as a good measure of "caseness" (i.e., the higher the score, the more likely there is an underlying depressive or anxiety disorder). In particular, the PHQ-2 and GAD-2 subscores of the PHQ-4 provide separate depressive and anxiety scores, and can be used as screeners for depression and anxiety.

PHQ-15 Somatic Symptom Severity. This is calculated by assigning scores of 0, 1, and 2 to the response categories of "not at all", "bothered a little", and "bothered a lot", for the 13 somatic symptoms of the PHQ (items 1a-1m). Also, 2 items from the depression module (sleep and tired) are scored 0 ("not at all"), 1 ("several days") or 2 ("more than half the days" or "nearly every day"). Thus, a PHQ-15 score can be derived from page 1 of the PHQ, or from separate administration of the PHQ-15 scale or the PHQ-SADS. PHQ-15 scores of 5, 10, and 15 represent cutpoints for low, medium, and high somatic symptom severity, respectively.

Sensitivity to Change for Monitoring Treatment Outcomes. A particularly important use of a measure is its responsiveness to changes of condition severity over time. This is well-established for the PHQ-9 which is increasingly used as a measure to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (to determine treatment response).[6,10] An example of how different PHQ-9 severity levels might guide treatment is shown in **Table 4, page 7**. There is preliminary evidence that the PHQ-15 may be responsive to changes as individuals with somatoform disorders or high somatization are treated.[10] The GAD-7 has demonstrated change as a secondary anxiety outcome in several depression trials, but has not yet been studied as a primary outcome in anxiety trials. Also, since there is more diagnostic splitting for anxiety than for depressive disorders, it remains to be determined whether a single anxiety measure can suffice as an outcome measure. It is likely the GAD-7 will be useful but not yet certain it will be sufficient.

Psychometrics. The psychometrics of the PHQ and its component scales are described in the validation articles for specific measures (see Selected References on page 9) and are summarized in a review article on the PHQ-9, GAD-7, and PHQ-15.[10]

Table 4. PHQ-9 Scores and Proposed Treatment Actions *

PHQ-9 Score	Depression Severity	Proposed Treatment Actions
0 – 4	None-minimal	None
5 – 9	Mild	Watchful waiting; repeat PHQ-9 at follow-up
10 – 14	Moderate	Treatment plan, considering counseling, follow-up and/or pharmacotherapy
15 – 19	Moderately Severe	Active treatment with pharmacotherapy and/or psychotherapy
20 – 27	Severe	Immediate initiation of pharmacotherapy and, if severe impairment or poor response to therapy, expedited referral to a mental health specialist for psychotherapy and/or collaborative management

* From Kroenke K, Spitzer RL, Psychiatric Annals 2002;32:509-521

TRANSLATIONS

There are numerous translations of the PHQ as well as the PHQ-9 and GAD-7 available in many languages, which are freely downloadable on the PHQ website (www.phqscreeners.com). The abbreviated versions of these measures – PHQ-8, PHQ-2, GAD-2, and PHQ-4 – can simply be derived from the translations by selecting the relevant items (see Table 1, page 3). The PHQ-15 can also be simply derived by selecting the 13 somatic items (1a-1m), plus the *sleep* and *tired* items (2c and 2c) from the PHQ translations.

Many of the translations have been developed by the MAPI Research Institute using an internationally accepted translation methodology. Thus, most of the translations are linguistically valid. However, unlike the English versions of the PHQ and GAD-7, few of the translations have been psychometrically validated against an independent structured psychiatric interview.

WEBSITE

Copies of the PHQ family of measures, including the GAD-7, are available at the website:

www.phqscreeners.com

Also, translations, a bibliography, an instruction manual, and other information is provided on this website.

QUESTIONS REGARDING DEVELOPMENT, ACKNOWLEDGMENTS AND USE

The PHQ family of measures (see Table 1, page 3), including abbreviated and alternative versions as well as the GAD-7, were developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke and colleagues, with an educational grant from Pfizer Inc.

All of the measures included in Table 1 are in the public domain. No permission is required to reproduce, translate, display or distribute.

SELECTED REFERENCES

1. Spitzer RL, Williams JBW, Kroenke K, Linzer M, deGruy FV, Hahn SR, Brody D, Johnson JG. Utility of a new procedure for diagnosing mental disorders in primary care: The PRIME-MD 1000 study. JAMA 1994;272:1749-1756.
2. Spitzer RL, Kroenke K, Williams JBW, for the Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. JAMA 1999;282:1737-1744.
3. Spitzer RL, Williams JBW, Kroenke K, et al. Validity and utility of the Patient Health Questionnaire in assessment of 3000 obstetrics-gynecologic patients. Am J Obstet Gynecol 2000; 183:759-769
4. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001;16:606-613.
5. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals 2002;32:509-521. *[also includes validation data on PHQ-8*]
6. Löwe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42:1194-1201
7. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166:1092-1097.
8. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007;146:317-325. *[validation data on GAD-7 and GAD-2 in detecting 4 common anxiety disorders)]*
9. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: Validity of a new measure for evaluating somatic symptom severity. Psychosom Med 2002;64:258-266.
10. Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010 (in press).
11. Johnson JG, Harris ES, Spitzer RL, Williams JBW. The Patient Health Questionnaire for Adolescents: Validation of an instrument for the assessment of mental disorders among adolescent primary care patients. J Adolescent Health. 2002;30:196-204.
12. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003; 41:1284-1292.
13. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50:613-621
14. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affective Disorders 2009;114:163-173.
15. Löwe B, Spitzer RL, Williams JBW, Mussell M, Schellberg D, Kroenke K. Depression, anxiety, and somatization in primary care: syndrome overlap and functional impairment. Gen Hosp Psychiatry 2008;30:191-199.
16. Dube P, Kroenke K, Bair MJ, Theobald D, Williams L. The P4 screener: a brief measure for assessing potential suicidal risk. J Clin Psychiatry Primary Care Companion 2010 (in press). *[Algorithm for following up on positive responses to 9th item of PHQ-9]*