Patient Health Questionnaire-9 (PHQ-9)

Handbook of Assessment in Mindfulness Research

Note: As a reference work, please avoid first-person usage in the writing of your contribution. Please refer to the Guidelines for Authors for more details.

Length of text: Word count in the range of (12000-15000) words.

PATIENT HEALTH QUESTIONNAIRE-9 (PHQ-9)

Author's name and details

Rebecca Y. M. Cheung*

School of Psychology and Clinical Language Sciences, University of Reading

*Corresponding author; Correspondence should be directed to Rebecca Y. M. Cheung, Ph.D., School of Psychology and Clinical Language Sciences, Harry Pitt Building, University of Reading, Earley Gate, Reading, RG6 7BE, United Kingdom. E-mail: rebecca.cheung@reading.ac.uk. ORCID: https://orcid.org/0000-0003-0998-7991

Abstract

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item instrument for assessing depression commonly used in mindfulness research. It is both a measure of depression severity and a diagnostic measure of depressive disorders. The PHQ-9 has been translated and validated in multiple languages and contexts involving clinical and population-based samples. In terms of factor-structure, a majority of existing studies recommend either a one-factor model of depression or a two-factor model with affective / cognitive (i.e., non-somatic) and somatic dimensions. The PHQ-9 is brief, freely available, easy to score, has diagnostic utility as well as good psychometric properties, and can be administered in various ways, such as paper-based self-administration, telephone, smartphone custom app, and chatbot. Shorter and validated versions are available as PHQ-8 and PHQ-2 for the purpose of research or screening. Overall, the PHQ-9 is a useful tool for assessing depression.

Keywords

Depression severity; depressive disorders; depressive symptoms; diagnostic instrument

Introduction

Over the last decade, research on the link between mindfulness and depression has flourished (e.g., Desrosiers et al., 2013; Petrocchi & Ottaviani, 2016). Studies reveal that mindfulness is associated with depression through processes such as disengagement from autopilot, reduction in ruminative thinking, and increases in present-focused attention, nonelaborative awareness, and nonjudgmental acceptance (Bishop et al., 2004; Brown & Ryan, 2003; Kabat-Zinn, 1982). Among existing measures of depression, the Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001; Kroenke & Spitzer, 2002) is commonly used in mindfulness research (e.g., Cheung et al., 2019; Querstret et al., 2018; Segal et al., 2020). The PHQ-9 is a 9-item measure based on the diagnostic criteria of Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV; American Psychiatric Association, 1994) for depression (Kroenke et al., 2001). The PHQ-9 is the depression module from the Patient Health Questionnaire (PHQ), which is a self-administrated version of the primary care evaluation of mental disorders (PRIME-MD®) designed to assess DSM-IV disorders including mood, anxiety, somatoform, eating, and alcohol abuse/dependence (Spitzer et al., 1994). The PHQ has been validated in two large-scale studies involving samples

of 3,000 adult patients from 9 primary care clinics (Spitzer et al., 1999) and another 3,000 adults from 7 obstetrics-gynecology clinics (Spitzer et al., 2000).

Main Text

With a total of 9 items, people rate each item of the PHQ-9 on how often they have experienced those specific symptoms of depression over the last two weeks on a scale of 0 (“not at all”), 1 (“several days”), 2 (“more than half the days”), and 3 (“nearly every day”). As a measure of severity, the scores are then summed to a score ranging from 0 to 27, with higher scores indicating a greater level of severity. An additional tenth item is included at the end of the measure as a global rating of functional impairment associated with the depressive symptoms. The cutoff points for scores are 1 to 4 (“none”), 5 to 9 (“mild”), 10 to 14 (“moderate”), 15 to 19 (“moderately severe”), and 20 to 27 (“severe”). A summed score of 10 or greater suggests a seven-fold likelihood of major depression, whereas a summed score of 15 or greater indicates the presence of depression (Kroenke et al., 2002). As a diagnostic measure of major depressive disorder, five of the nine symptom criteria must be present for “more than half the days” over the past two weeks, with one of the symptoms being anhedonia or depressed mood. Disregarding the duration, the symptom “thoughts that you would be better off dead or of hurting yourself in some way” counts towards the nine symptoms. As a diagnostic measure of depression other than major depressive disorder, two to four symptoms must be present for “more than half the days” over the past two weeks, with one of the symptoms being anhedonia or depressed mood. Even though the DSM has been updated to the fifth version as the DSM-5 (American Psychiatric Association, 2013), the nine depressive symptom criteria as well as the two-week period of diagnostic symptoms remain the same. Hence, the PHQ-9 remains a relevant instrument of depression to-date.

The PHQ-9 was initially validated in primary care patients (Kroenke et al., 2001) and demonstrated good internal consistency (Cronbach’s alpha = .89) and test-retest reliability (e.g., correlation between test-retest scores within 48 hours = .84). In the Kroenke et al.’s study (2001), criterion validity was also demonstrated. Specifically, greater PHQ-9 scores were related to a greater likelihood of major depression in agreement with the diagnoses made by mental health professionals. The area under the ROC curve was .95, thereby suggesting that the PHQ-9 can distinguish between people with and without major depression. In terms of construct validity, the PHQ-9 was correlated with lower functional status, such as mental, social, and role functioning, as measured by the Medical Outcomes Study Short-Form General Health Survey (SF-20; Stewart et al., 1988).

Subsequent Evidence of Psychometric Properties

The PHQ-9 was subsequently validated in other samples, such as ethnically diverse primary care patients (Huang et al., 2006), psychiatric patients (Beard et al., 2016), HIV-infected patients (Crane et al., 2010), college students (Granillo, 2012; Keum et al., 2018), and a population-based sample in the United States (Thibodeau & Asmundson, 2014). It was also validated in samples of cancer patients (Hinz et al., 2016), medical outpatients with major depressive disorder, depressive disorders, or no depressive disorders (Löwe et al., 2004), an elderly general population-based sample (Forkmann et al., 2013), and a population-based sample in Germany (Martin et al., 2006). Other than the United States and Germany, the PHQ-9 was validated in diverse samples worldwide, such as a multi-ethnic population-based sample in Europe (Galenkamp et al., 2017), Chinese adolescents and a population-based sample in Hong Kong (Leung et al., 2019; Yu et al., 2012), adults living with HIV/AIDS in western Kenya (Monahan et al., 2009), adults with autism in Australia (Arnold et al., 2019), people with visual impairment in South India (Gothwal et al., 2014), Filipina and Indonesian female migrant domestic workers in Macao (Hall et al., 2021), ambulatory care patients in Argentina (Urtasun et al., 2019), pregnant women in Spain (Marcos-Nájera et

al., 2018), patients receiving palliative care services in the United Kingdom (Chilcot et al., 2013), and diabetes patients in the Netherlands (van Steenbergen-Weijenburg et al., 2010).

In the subsequent validations, differential item functioning (DIF) analyses generally indicated that the PHQ-9 assesses depression similarly in men and women (Thibodeau & Asmundson, 2014) and in diverse ethnic groups (Galenkamp et al., 2017). However, findings based on smaller and specific samples did show race-, age-, or gender-related DIF on various items (e.g., Crane et al., 2010; Forkmann et al., 2013; Huang et al., 2006). Findings based on exploratory and confirmatory factor analyses are also mixed. Notably, some studies indicated a one-factor model for population-based samples, elderly general population, Hispanic LGBT+ community, primary care patients, and racially diverse college students (e.g., González-Rivera, 2019; Forkmann et al., 2013; Galenkamp et al., 2017; Huang et al., 2006; Keum et al., 2018; Kocalevent et al., 2013; Yu et al., 2012; Thibodeau & Asmundson, 2014). Others indicated a bifactor / two-factor model with affective / cognitive or psychological (i.e., non-somatic) and somatic dimensions of depression emerging for psychiatric patients, patients receiving palliative care service, cancer patients, Latina and Non-Latina white female college students, female migrant domestic workers, National Guard soldiers, and people of white British ancestry in the United Kingdom (e.g., Beard et al., 2016; Chilcot et al., 2013; Elhai et al., 2012; Granillo, 2012; Hall et al., 2021; Hinz et al., 2016; Thorp et al., 2019). Furthermore, an adult twin study conducted in the United States suggested a three-factor model comprising factors of cognitive / motor, mood, and neurovegetative symptoms (Kendler et al., 2013). Another study also indicated a third factor of pregnancy-related symptoms, in addition to factors of cognitive-affective and somatic symptoms, among pregnant women such as changes in appetite (Marcos-Nájera et al., 2018). Therefore, depending on the analytic approach or the population of interest, the PHQ-9 shows different psychometric properties, ranging from one-, two-, or three factor models to psychometric equivalence vs. nonequivalence of individual items between various groups (e.g., race, gender). Regardless of these differences in the psychometric properties, construct validity and criterion validity were established for the PHQ-9 against diagnostic interviews and other instruments of health or depression (e.g., Leung et al., 2019; Kocalevent et al., 2013; Urtasun et al., 2019).

Short Forms: Validation and Psychometric Properties

Although the PHQ-9 is relatively brief (Kroenke et al., 2001), even shorter versions have been created and validated, including the PHQ-8 (Kroenke et al., 2009) and the PHQ-2 (Kroenke et al., 2003). As a measure of diagnosis and severity for depressive disorders, the PHQ-8 (Kroenke et al., 2009) has eight of the nine criteria of depressive disorders based on the DSM-IV (American Psychiatric Association, 1994), with the exclusion of the ninth criterion aiming to assess self-injurious or suicidal thoughts, given that the item is often endorsed the least frequently on the PHQ-9 and removing it minimally affects the overall scoring (Kroenke & Spitzer, 2002). Kroenke and Spitzer (2002) further recommended that the PHQ-8 may be used in clinical samples or population samples meeting one of the three criteria: (a) low or negligible risk of suicidality, (b) assessment of depression as a secondary outcome in studies about other medical conditions, and (c) data collection through self-administration instead of direct interview (see also Kroenke et al., 2009). The rating scale of the PHQ-8 is similar to that of the PHQ-9, with cutoff points at 1 to 4 (“none”), 5 to 9 (“mild”), 10 to 14 (“moderate”), 15 to 19 (“moderately severe”), and 20 to 24 (“severe”). The psychometric properties of PHQ-8 are also similar to those of the PHQ-9. Persons with a depressive disorder are 15 times more likely to have a summed score of 10-14 on the PHQ-8 than persons without a depressive disorder (Kroenke et al., 2009). Subsequent findings based on confirmatory factor analyses are mixed, with some studies showing a one-factor model, e.g., among Puerto Rican adults, Mexican and Central American descent university students, and Hispanic LGBT+ community (Alpizar et al., 2018; González-Rivera, 2019; Pagán-Torres et al., 2020), whereas others showing a two-factor model with affective / cognitive and somatic dimensions of depression, e.g., among persons with systemic sclerosis in Sweden and primary care patients in

Germany (Mattsson et al., 2020; Moehring et al., 2021). Hence, the factor structure of the PHQ-8 varies in different populations.

The PHQ-2 (Kroenke et al., 2003) assesses the frequency of depressed mood and anhedonia (i.e., the first two items of the PHQ-9) on a scale of 0 ("not at all") to 3 ("nearly every day") over the past two weeks. Although Kroenke et al. (2003) recommended the PHQ-9 over the PHQ-2 either as a diagnostic instrument or an assessment of severity, the PHQ-2 may be useful as a screening measure for depression. The PHQ-2 demonstrated criterion and construct validity (Kroenke et al., 2003). When assessed against a mental health professional interview, a score of 3 or above indicated high sensitivity and specificity for major depressive disorder. The optimal cutoff point of PHQ-2 is 3 for the purpose of screening.

Scale Versions in Other Languages: Validation and Psychometric Properties

The PHQ-9 has been translated and validated in multiple languages and contexts, including Arabic (Sawaya et al., 2016), Argentinian (Urtasun et al., 2019), Chinese (Chen et al., 2010; Leung et al., 2019), Dutch (van Steenbergen-Weijenburg et al., 2010), French (Carballeira et al., 2007) and Canadian French (Arthurs et al., 2012), German (Martin et al., 2006; Reich et al, 2018), Japanese (Muramatsu et al., 2018), Korean (Han et al., 2008), Latvian and Russian (Rancans et al., 2018), Norwegian (Wisting et al., 2021), Portuguese (Lamela et al., 2020), Spanish (Muñoz-Navarro et al., 2017; Marcos-Nájera et al., 2018), Swahili (Fawzi et al., 2019), Thai (Lotrakul et al., 2008), and Turkish (Reich et al, 2018). Once again, findings based on factor analyses were mixed, with some studies showing a one-factor model (e.g., Leung et al., 2019), others showing a two-factor model with affective / cognitive and somatic dimensions of depression (e.g., Sawaya et al., 2016), and one study showing that both one- and two-factor models were adequate (Arthurs et al., 2012). Regardless of the differences in the psychometric properties across translations, construct validity and criterion validity were established for the PHQ-9 against diagnostic interviews and other instruments of health or depression (e.g., Carballeira et al., 2007; Muramatsu et al., 2018; Urtasun et al., 2019).

Ways of Administration

The PHQ-9 is originally developed as a self-administered measure (Kroenke & Spitzer, 2002). People rate each of the 9 items on how often they have experienced symptoms of depression during the last two weeks (i.e., "Over the last 2 weeks, how often have you been bothered by any of the following problems?") on a scale of 0 ("not at all"), 1 ("several days"), 2 ("more than half the days"), and 3 ("nearly every day"). A tenth item is included at the end of the measure as a global rating of functional impairment associated with the depressive symptoms. The tenth item reads, "If you checked off 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?" People rate on a scale of "not difficult at all", "somewhat difficult", "very difficult", and "extremely difficult." A copy of the scale with instructions can be found in the Appendix of Kroenke et al. (2001, p. 613).

Aside from paper-based self-administration, the PHQ-9 has also been administered over the telephone (Pinto-Meza et al., 2005), by smartphone custom app (Torous et al., 2015), interview (Wulsin et al., 2002), and chatbot (Dosovitsky et al., 2021). Findings suggested that the alternative ways of administration were feasible and that the psychometric properties were adequate.

Scoring the Data

The scoring of the PHQ-9 depends on whether a one-factor model or a two-factor model is adopted. For one-factor models, the raw scores of items 1 to 9 are summed to a score ranging from 0 to 27, with higher scores indicating a greater severity of depression, as recommended by Kroenke and Spitzer (2012). For two-factor models, different scorings for the cognitive-affective factor and the somatic factor have been proposed in previous

research (e.g., Elhai et al., 2012; Keum et al., 2018). For example, Elhai et al (2012) found a two-factor model with items 1, 2, 6, and 9 comprising the cognitive-affective factor and items 3, 4, 5, 7, and 8 the somatic factor. The items can be summed, with higher scores indicating greater severity of the respective factors of depression.

Strengths and Limitations

Two decades of research has advanced our understanding of depression via the development, validation, and utility of the PHQ-9. The PHQ-9 has numerous strengths including its brevity, free availability, diagnostic utility, good psychometric properties, and procedural validity via paper-based self-administration, telephone (Pinto-Meza et al., 2005), smartphone custom app (Torous et al., 2015), interview (Wulsin et al., 2002), and chatbot (Dosovitsky et al., 2021). The measure should, however, be considered in light of several limitations. For instance, although both a unidimensional one-factor model or a two-factor model are commonly reported in the literature (e.g., Beard et al., 2016; Galenkamp et al., 2017), inconsistencies between the factor structure of the PHQ-9 may lead to problems with scoring and conceptualization. Likewise, the inconsistences between DIF for specific items (e.g., Crane et al., 2010; Forkmann et al., 2013; Huang et al., 2006) and the null DIF findings (e.g., Galenkamp et al., 2017; Thibodeau & Asmundson, 2014) may require further clarification and validation. Finally, despite its diagnostic utility, assessment by mental health professional interview may be necessary to confirm any potential clinical diagnoses.

Conclusion

The PHQ-9 is an instrument for detecting depression commonly used in mindfulness research (e.g., Cheung et al., 2019; Querstret et al., 2018; Segal et al., 2020). It is a 9-item measure of depression severity and a diagnostic measure of depressive disorders (Kroenke et al., 2001). The PHQ-9 has been translated and validated in multiple languages and contexts involving clinical and population-based samples. It is brief, freely available, easy to score, has diagnostic utility as well as good psychometric properties, and can be administered in various ways, such as paper-based self-administration, telephone, smartphone custom app, and chatbot - although some confusion remains about the precise nature of its factor structure. Shorter versions are available as the 8-item PHQ-8 (Kroenke et al., 2009) and the 2-item PHQ-2 (Kroenke et al., 2003) for the purpose of research and screening. Overall, the PHQ-9 is a useful tool in assessing depression.

References

Alpizar, D., Lagana, L., Plunkett, S. W., & French, B. F. (2018). Evaluating the 8-item Patient Health Questionnaire's psychometric properties with Mexican and Central American descent university students. *Psychological Assessment*, *30*(6), 719–728. https://doi.org/10.1037/pas0000521

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

González-Rivera, J. A. (2019). Validation and dimensionality of Patient Health Questionnaire for depression (PHQ-8 and PHQ-9) in Hispanic LGBT+ community. *International Journal of Recent Scientific Research*. *10*(12), 36670–36676. http://doi.org/10.24327/ijrsr.2020.1012.4970

Arnold, S. R., Uljarević, M., Hwang, Y. I., Richdale, A. L., Trollor, J. N., & Lawson, L. P. (2019). Brief report: Psychometric properties of the Patient Health Questionaire-9 (PHQ-9) in autistic adults. *Journal of Autism and Developmental Disorders*, *50*(6), 2217–2225. https://doi.org/10.1007/s10803-019-03947-9

Arthurs, E., Steele, R. J., Hudson, M., Baron, M., & Thombs, B. D. (2012). Are scores on English and French versions of the PHQ-9 comparable? An assessment of differential item functioning. *PloS One*, *7*(12), e52028. https://doi.org/10.1371/journal.pone.0052028

Handbook of Assessment in Mindfulness Research

Note: As a reference work, please avoid first-person usage in the writing of your contribution. Please refer to the Guidelines for Authors for more details.

Length of text: Word count in the range of (12000-15000) words.

Beard, C., Hsu, K., Rifkin, L., Busch, A., & Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*, *193*, 267–273. https://doi.org/10.1016/j.jad.2015.12.075

Bishop, S. M., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., Segal, Z. V., Abbey, S., Speca, M., Velting, D., & Devins, G. (2004). Mindfulness: A proposed operational definition. *Clinical Psychology: Science and Practice, 11*(3), 230–241. https://doi.org/10.1093/clipsy.bph077

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present. *Journal of Personality and Social Psychology*, *84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Carballeira, Y., Dumont, P., Borgacci, S., Rentsch, D., de 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*, *80*(1), 69–77. https://doi.org/10.1348/147608306X103641

Chen, S., Chiu, H., Xu, B., Ma, Y., Jin, T., Wu, M., & Conwell, Y. (2010). Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. *International Journal of Geriatric Psychiatry*, *25*(11), 1127–1133. https://doi.org/10.1002/gps.2442

Cheung, R. Y. M., & Ng, M. C. Y. (2019). Mindfulness and symptoms of depression and anxiety: The underlying roles of awareness, acceptance, impulse control, and emotion regulation. *Mindfulness*, *10*(6), 1124–1135. https://doi.org/10.1007/s12671-018-1069-y

Chilcot, J., Rayner, L., Lee, W., Price, A., Goodwin, L., Monroe, B., Sykes, N., Hansford, P., & Hotopf, M. (2013). The factor structure of the PHQ-9 in palliative care. *Journal of Psychosomatic Research*, *75*(1), 60–64. https://doi.org/10.1016/j.jpsychores.2012.12.012

Crane, P. K., Gibbons, L. E., Willig, J. H., Mugavero, M. J., Lawrence, S. T., Schumacher, J. E., Saag, M. S., Kitahata, M. M., & Crane, H. M. (2010). Measuring depression levels in HIV-infected patients as part of routine clinical care using the nine-item Patient Health Questionnaire (PHQ-9). *AIDS Care*, *22*(7), 874–885. https://doi.org/10.1080/09540120903483034

Desrosiers, A., Klemanski, D. H., & Nolen-Hoeksema, S. (2013). Mapping mindfulness facets onto dimensions of anxiety and depression. *Behavior Therapy*, *44*(3), 373–384. https://doi.org/10.1016/j.beth.2013.02.001\

Dosovitsky, G., Kim, E., & Bunge, E. L. (2021). Psychometric properties of a chatbot version of the PHQ-9 with adults and older adults. *Frontiers in Digital Health*, *3*, 645805. https://doi.org/10.3389/fdgth.2021.645805

Elhai, J. D., Contractor, A. A., Tamburrino, M., Fine, T. H., Prescott, M. R., Shirley, E., Chan, P. K., Slembarski, R., Liberzon, I., Galea, S., & Calabrese, J. R. (2012). The factor structure of major depression symptoms: A test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Research*, *199*(3), 169-173. https://doi.org/10.1016/j.psychres.2012.05.018

Fawzi, M. C. S., Ngakongwa, F., Liu, Y., Rutayuga, T., Siril, H., Somba, M., & Kaaya, S. (2019). Validating the Patient Health Questionnaire-9 (PHQ-9) for screening of depression in Tanzania. *Neurology, Psychiatry, and Brain Research*, *31*, 9-14. https://doi.org/10.1016/j.npbr.2018.11.002

Forkmann, T., Gauggel, S., Spangenberg, L., Brähler, E., & Glaesmer, H. (2013). Dimensional assessment of depressive severity in the elderly general population: Psychometric evaluation of the PHQ-9 using Rasch analysis. *Journal of Affective Disorders*, *148*(2-3), 323–330. https://doi.org/10.1016/j.jad.2012.12.019

Galenkamp, H., Stronks, K., Snijder, M. B., & Derks, E. M. (2017). Measurement invariance testing of the PHQ-9 in a multi-ethnic population in Europe: The HELIUS study. *BMC Psychiatry*, *17*(1), 349. https://doi.org/10.1186/s12888-017-1506-9

Gothwal, V. K., Bagga, D. K., & Sumalini, R. (2014). Rasch validation of the PHQ-9 in people with visual impairment in South India. *Journal of Affective Disorders*, *167*, 171–177. https://doi.org/10.1016/j.jad.2014.06.019

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. https://doi.org/10.5243/jsswr.2012.6

Hall, B. J., Patel, A., Lao, L., Liem, A., Mayawati, E. H., & Tjipto, S. (2021). Structural validation of the Patient Health Questionnaire-9 (PHQ-9) among Filipina and Indonesian female migrant domestic workers in Macao: Structural validation of PHQ-9. *Psychiatry Research*, *295*, 113575. https://doi.org/10.1016/j.psychres.2020.113575

Handbook of Assessment in Mindfulness Research

Note: As a reference work, please avoid first-person usage in the writing of your contribution. Please refer to the Guidelines for Authors for more details.

Length of text: Word count in the range of (12000-15000) words.

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

Hinz, A., Mehnert, A., Kocalevent, R. D., Brähler, E., Forkmann, T., Singer, S., & Schulte, T. (2016). Assessment of depression severity with the PHQ-9 in cancer patients and in the general population. *BMC Psychiatry*, *16*(1), 22. https://doi.org/10.1186/s12888-016-0728-6

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 Internal Medicine: JGIM*, *21*(6), 547–552. https://doi.org/10.1111/j.1525-1497.2006.00409.x

Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry*, *4*(1), 33–47. https://doi.org/10.1016/0163-8343(82)90026-3

Kendler, K. S., Aggen, S. H., & Neale, M. C. (2013). Evidence for multiple genetic factors underlying DSM-IV criteria for major depression. *JAMA Psychiatry, 70*(6), 599-607. https://doi.org/10.1001/jamapsychiatry.2013.751

Keum, B. T., Miller, M. J., & Inkelas, K. K. (2018). Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse U.S. college students. *Psychological Assessment*, *30*(8), 1096–1106. https://doi.org/10.1037/pas0000550

Kocalevent, R. D., Hinz, A., & Brähler, E. (2013). Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry*, *35*(5), 551–555. http://doi.org/10.1016/j.jad.2012.12.019

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, *32*(9), 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine: JGIM*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, *41*(11), 1284–1292. https://doi.org/10.1097/01.MLR.0000093487.78664.3C

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*(1), 163–173. https://doi.org/10.1016/j.jad.2008.06.026

Lamela, D., Soreira, C., Matos, P., & Morais, A. (2020). Systematic review of the factor structure and measurement invariance of the Patient Health Questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings. *Journal of Affective Disorders*, *276*, 220–233. https://doi.org/10.1016/j.jad.2020.06.066

Leung, D. Y. P., Mak, Y. W., Leung, S. F., Chiang, V. C. L., & Loke, A. Y. (2020). Measurement invariances of the PHQ-9 across gender and age groups in Chinese adolescents. *Asia-Pacific Psychiatry*, *12*(3). https://doi.org/10.1111/appy.12381

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

Löwe, B., Kroenke, K., Herzog, W., & Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders*, *81*(1), 61–66. https://doi.org/10.1016/s0165-0327(03)00198-8

Marcos-Nájera, R., Le, H. N., Rodríguez-Muñoz, M. F., Olivares Crespo, M. E., & Izquierdo Mendez, N. (2018). The structure of the Patient Health questionnaire-9 in pregnant women in Spain. *Midwifery*, *62*, 36–41. https://doi.org/10.1016/j.midw.2018.03.011

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry, 28*(1), 71–77. https://doi.org/10.1016/j.genhosppsych.2005.07.003

Mattsson, M., Sandqvist, G., Hesselstrand, R., Nordin, A., & Boström, C. (2020). Validity and reliability of the Patient Health Questionnaire-8 in Swedish for individuals with systemic sclerosis. *Rheumatology International*, *40*(10), 1675–1687. https://doi.org/10.1007/s00296-020-04641-1

Handbook of Assessment in Mindfulness Research

Note: As a reference work, please avoid first-person usage in the writing of your contribution. Please refer to the Guidelines for Authors for more details.

Length of text: Word count in the range of (12000-15000) words.

Moehring, A., Guertler, D., Krause, K., Bischof, G., Rumpf, H. J., Batra, A., Wurm, S., John, U., & Meyer, C. (2021). Longitudinal measurement invariance of the Patient Health Questionnaire in a German sample. *BMC Psychiatry*, *21*(1), 386. https://doi.org/10.1186/s12888-021-03390-0

Monahan, P. O., Shacham, E., Reece, M., Kroenke, K., Ong'or, W. O., Omollo, O., Yebei, V. N., & Ojwang, C. (2008). Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in Western Kenya. *Journal of General Internal Medicine: JGIM*, *24*(2), 189–197. https://doi.org/10.1007/s11606-008-0846-z

Muñoz-Navarro, R., Cano-Vindel, A., Medrano, L. A., Schmitz, F., Ruiz-Rodríguez, P., Abellán-Maeso, C., Font-Payeras, M. A., & Hermosilla-Pasamar, A. M. (2017). Utility of the PHQ-9 to identify major depressive disorder in adult patients in Spanish primary care centres. *BMC Psychiatry*, *17*(1), 291. https://doi.org/10.1186/s12888-017-1450-8

Muramatsu, K., Miyaoka, H., Kamijima, K., Muramatsu, Y., Tanaka, Y., Hosaka, M., Miwa, Y., Fuse, K., Yoshimine, F., Mashima, I., Shimizu, N., Ito, H., & Shimizu, E. (2018). Performance of the Japanese version of the Patient Health Questionnaire-9 (J-PHQ-9) for depression in primary care. *General Hospital Psychiatry*, *52*, 64-69. https://doi.org/10.1016/j.genhosppsych.2018.03.007

Pagán-Torres, O. M., González-Rivera, J. A., & Rosario-Hernández, E. (2020). Psychometric analysis and factor structure of the Spanish version of the eight-item Patient Health Questionnaire in a general sample of Puerto Rican Adults. *Hispanic Journal of Behavioral Sciences*, *42*(3), 401–415. https://doi.org/10.1177/0739986320926524

Petrocchi, N., & Ottaviani, C. (2016). Mindfulness facets distinctively predict depressive symptoms after two years: The mediating role of rumination. *Personality and Individual Differences*, *93*, 92–96. https://doi.org/10.1016/j.paid.2015.08.017

Pinto-Meza, A., Serrano-Blanco, A., Peñarrubia, 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*(8), 738–742. https://doi.org/10.1111/j.1525-1497.2005.0144.x

Querstret, D., Cropley, M., & Fife-Schaw, C. (2018). The effects of an online mindfulness intervention on perceived stress, depression and anxiety in a non-clinical sample: A randomised waitlist control trial. *Mindfulness*, *9*(6), 1825–1836. https://doi.org/10.1007/s12671-018-0925-0

Rancans, E., Trapencieris, M., Ivanovs, R., & Vrublevska, J. (2018). Validity of the PHQ-9 and PHQ-2 to screen for depression in nationwide primary care population in Latvia. *Annals of General Psychiatry*, *17*(1), 33. https://doi.org/10.1186/s12991-018-0203-5

Reich, H., Rief, W., Brähler, E., & Mewes, R. (2018). Cross-cultural validation of the German and Turkish versions of the PHQ-9: An IRT approach. *BMC Psychology*, *6*(1), 26–26. https://doi.org/10.1186/s40359-018-0238-z

Sawaya, H., Atoui, M., Hamadeh, A., Zeinoun, P., & Nahas, Z. (2016). Adaptation and initial validation of the Patient Health Questionnaire – 9 (PHQ-9) and the Generalized Anxiety Disorder – 7 Questionnaire (GAD-7) in an Arabic speaking Lebanese psychiatric outpatient sample. *Psychiatry Research*, *239*, 245–252. https://doi.org/10.1016/j.psychres.2016.03.030

Segal, Z. V., Dimidjian, S., Beck, A., Boggs, J. M., Vanderkruik, R., Metcalf, C. A., Gallop, R., Felder, J. N., & Levy, J. (2020). Outcomes of online mindfulness-based cognitive therapy for patients with residual depressive symptoms: A randomized clinical trial. *JAMA Psychiatry*, *77*(6), 563–573. https://doi.org/10.1001/jamapsychiatry.2019.4693

Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. *JAMA*, *282*(18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737

Spitzer, R. L., Williams, J. B., Kroenke, K., Hornyak, R., & McMurray, J. (2000). Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire obstetrics-gynecology study. *American Journal of Obstetrics and Gynecology*, *183*(3), 759–769. https://doi.org/10.1067/mob.2000.106580

Spitzer, R. L., Williams, J. B. W., Kroenke, K., Linzer, M., deGruy, F. V. III, Hahn, S. R., Brody, D., & Johnson, J. G. (1994). Utility of a new procedure for diagnosing mental disorders in primary care: The PRIME-MD 1000 study. *JAMA*, *272*(22), 1749–1756. https://doi.org/10.1001/jama.1994.03520220043029

Stewart, A. L., Hays, R. D., & Ware, J. E. (1988). The MOS short-form general health survey: Reliability and validity in a patient population. *Medical Care*, *26*(7), 724–735. https://doi.org/10.1097/00005650-198807000-00007

Handbook of Assessment in Mindfulness Research

Note: As a reference work, please avoid first-person usage in the writing of your contribution. Please refer to the Guidelines for Authors for more details.

Length of text: Word count in the range of (12000-15000) words.

Thibodeau, M. A., & Asmundson, G. J. (2014). The PHQ-9 assesses depression similarly in men and women from the general population. *Personality and Individual Differences*, *56*(1), 149–153. https://doi.org/10.1016/j.paid.2013.08.039

Thorp, J. G., Marees, A. T., Ong, J. S., An, J., MacGregor, S., & Derks, E. M. (2020). Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9. *Psychological Medicine, 50*(14), 2385-2396. https://doi.org/10.1017/S0033291719002526

Torous, J., Staples, P., Shanahan, M., Lin, C., Peck, P., Keshavan, M., & Onnela, J. P. (2015). Utilizing a personal smartphone custom app to assess the Patient Health Questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Mental Health*, *2*(1), e8. https://doi.org/10.2196/mental.3889

Urtasun, M., Daray, F. M., Teti, G. L., Coppolillo, F., Herlax, G., Saba, G., Rubinstein, A., Araya, R., & Irazola, V. (2019). Validation and calibration of the Patient Health Questionnaire (PHQ-9) in Argentina. *BMC Psychiatry*, *19*(1), 291. https://doi.org/10.1186/s12888-019-2262-9

van Steenbergen-Weijenburg, K. M., de Vroege, L., Ploeger, R. R., Brals, J. W., Vloedbeld, M. G., Veneman, T. F., Hakkaart-van Roijen, L., Rutten, F. F. H., Beekman, A. T. F., & van der Feltz-Cornelis, C. M. (2010). Validation of the PHQ-9 as a screening instrument for depression in diabetes patients in specialized outpatient clinics. *BMC Health Services Research*, *10*(1). https://doi.org/10.1186/1472-6963-10-235

Wisting, L., Johnson, S. U., Bulik, C. M., Andreassen, O. A., Øyvind, R., & Bang, L. (2021). Psychometric properties of the Norwegian version of the Patient Health Questionnaire-9 (PHQ-9) in a large female sample of adults with and without eating disorders. *BMC Psychiatry*, *21*(1), 6. https://doi.org/10.1186/s12888-020-03013-0

Wulsin, L., Somoza, E., & Heck, J. (2002). The feasibility of using the Spanish PHQ-9 to screen for depression in primary care in Honduras. *Primary Care Companion to the Journal of Clinical Psychiatry*, *4*(5), 191–195. https://doi.org/10.4088/PCC.v04n0504

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. https://doi.org/10.1016/j.comppsych.2010.11.002