



Published in final edited form as:

Am J Geriatr Psychiatry. 2010 December ; 18(12): 1116–1123. doi:10.1097/JGP.0b013e3181dd1c26.

SCREENING FOR DEPRESSION IN AGING SERVICES CLIENTS

Thomas M. Richardson, PhD, MBA¹, Hua He, PhD², Carol Podgorski, PhD¹, Xin Tu, PhD^{1,2}, and Yeates Conwell, MD¹

¹University of Rochester School of Medicine and Dentistry, Department of Psychiatry

²University of Rochester School of Medicine and Dentistry, Department of Biostatistics

Abstract

Objectives—To establish the psychometric characteristics of the Patient Health Questionnaire (PHQ- 2, PHQ-9, and their sequential administration) in older adults who utilize community-based, social service care management.

Design—Comparison of screening tools with criterion standard diagnostic interview.

Setting—A community-based aging services agency.

Participants—378 adults age \geq 60 years undergoing in-home aging services care management assessments.

Measurements—Subjects were administered the PHQ-9 and Structured Clinical Interview for DSM-IV. We examined the sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and ROC curves separately for the PHQ-2 and the PHQ-9, and for a two stage screening process that used each in sequence (the PHQ-2/9).

Results—Using a cut-score of 3, the sensitivity of the PHQ-2 was .80 and the specificity was .78. The area under the ROC curve (AUC) for the PHQ-2 was .87. Using a cut-score of 10, the sensitivity and specificity of the PHQ-9 were .82 and .87. The AUC was .91. The sensitivity and specificity of the two-stage PHQ-2/9 were 0.81 and 0.89, and the AUC was 0.91.

Conclusions—The PHQ-9's greater specificity is an advantage over the PHQ-2 in aging service settings where false positive tests have potentially high cost. The PHQ-2/9 performed as well as the PHQ-9, but would be more efficient for the agency to administer. Combined with an appropriate referral system to healthcare providers, use of the PHQ-2/9 sequence by aging services personnel can efficiently assist in reducing the burden of late life depression.

Keywords

care management; psychometrics; social services; home and community-based care; mental health services

OBJECTIVE

The public health burden of affective illnesses is great. Depression is expected to be the second leading cause of disability adjusted life years (DALY's) worldwide by 2020 (1), and

(Corresponding Author), University of Rochester Medical Center, 300 Crittenden Blvd., Rochester, NY 14642, Phone: 585-275-6739, Fax: 585-273-1066, yeates_conwell@urmc.rochester.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

in the United States its associated direct and indirect costs have been estimated at over \$83 billion per year (2).

Depression in later life is common and disproportionately affects those with high medical burden and disability. Studies show that from 8% to 16% of community-dwelling older adults have clinically significant depressive symptoms, with higher rates among elderly women, the oldest old, and those with physical disability and cognitive impairment (3). In primary care, medical/surgical inpatient and long term care settings, up to a 35% of elderly patients have major, minor, or subsyndromal depression (4–6). Depression is associated with decreased quality of life (7), is a risk factor for suicide (8), and is independently associated with non-suicide mortality (9). Because subthreshold depression also results in many of the same untoward outcomes as major depression, it too is a target for prevention efforts (10,11).

Accurate and timely detection of major and subsyndromal depression is critical to reducing this burden, but depression among older adults is particularly challenging to identify and manage. Barriers to recognizing and effectively treating late life depression have been identified at the system (e.g., cost, continuity of care), provider (e.g., lack of training in geriatric care, competing demands on time), and patient levels (e.g., stigma, clinical complexity of resulting from comorbid conditions) (12,13). Despite the US Preventive Services Task Force recommending routine depression screening in primary care (14), these barriers limit the number of older adults who are identified and adequately treated for affective illness. Thus, depression screening in other venues where high-risk populations exist is warranted. One setting that may offer a unique and important opportunity for elders at risk for affective illness and its sequelae is the Aging Services Network (ASN).

The Older Americans Act of 1965 established the ASN as a system now comprising 56 State Units on Aging, 629 Area Agencies on Aging, 246 Tribal organizations, and nearly 20,000 community services provider organizations. Its mission is to help older adults maintain their independence at home by providing a range of services from information and referral programs to comprehensive care management, both in the seniors' homes and in community centers. The ASN serves over 7 million older adults and 300,000 caregivers each year (15), a high proportion of whom are at risk for depression due to social stressors, medical illness, and functional impairments (3). Studies that have examined the prevalence of mental disorders among ASN agency clients indicate that rates of clinically significant depression are high (12–33%)(16–18), but systematic efforts to detect and treat affective illness in community-based social service settings are rare (19). Broader dissemination of a depression screening tool well suited for use in the ASN would help address this disparity.

The effectiveness of a screening tool for depression requires that it be brief, easy to administer and interpret by a variety of providers in a variety of settings, and have good predictive properties (20). Several instruments have been developed to assist in the task (21). Perhaps the most widely disseminated in primary care settings is the Patient Health Questionnaire-9 (PHQ-9) (22,23), a valid and reliable screening tool for depressive syndromes among adults, including the elderly (24,25). Its nine questions correspond to the criteria set for major depression designated by the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) (26), allowing application of a DSM-IV-TR diagnostic algorithm for depressive disorders or a continuous measure of depressive symptom severity. The Patient Health Questionnaire-2 (PHQ-2) (27,28), an abbreviated version of the PHQ-9, has also been found to have similar properties among older adults and among Veterans (25,29,30). While not establishing a diagnosis of affective disorder, the PHQ-2 and the PHQ-9 are accurate in identifying persons who may benefit from a clinical assessment to determine if the illness is present. The utility of these instruments has been assessed

primarily in medical practice. The objective of this study was to examine their performance when administered to aging services care management clients, with the goal of providing guidance to social service agencies on what strategy for depression screening may be best suited to their particular service delivery system. Our specific aims were to examine the criterion validity (sensitivity, specificity, negative and positive predictive values and receiver operating curves) of the (a) PHQ-2, (b) the PHQ-9, and (c) the sequential administration of the PHQ-2 followed by the PHQ-9, in a sample of older community dwelling adults who received an in-home social work assessment for non-medical care management services.

METHODS

Setting

This study was conducted as part of a community-academic research partnership between a regional aging services provider, Eldersource (31), and the University of Rochester Medical Center. Eldersource is a social services agency certified by the Council on Accreditation (COA) (32), an accrediting body for human services agencies, to provide a range of non-medical services to help seniors and their families achieve or maintain optimal social, psychological, and physical functioning in the community. It serves as the point-of-entry for aging services in the Monroe County area, fielding over 13,000 requests for information and services each year. Those callers determined on initial telephone-based assessment to present more complex issues for which either short or longer-term care management may be indicated, are referred for in-home assessment.

Sample

Clients at least 60 years of age who spoke English and received an initial home assessment for Eldersource care management services from September 2005 to August 2007 were eligible to participate in the research interview. During the study period care managers referred 643 clients to the study, of which 378 clients provided informed consent and were enrolled (59%). Subjects (N=378) were primarily white (n=319; 84.4%) and female (n=259; 68.5%), 149 (39.4%) were married, and 234 (61.9%) had household incomes under \$1,750/month (<150% of the New York State poverty level for a family of two). Two hundred sixty-five (70.1%) had 12 years or more of education. The mean (\pm SD) age was 76.5 (\pm 9.2) years and nearly all subjects (n=368; 97.6%) reported having a primary care physician. One subject who did not have a PHQ-score was excluded from the analyses.

Care managers briefly introduced the study and, with verbal informed consent, referred interested clients to research personnel who contacted them, confirmed their eligibility, and invited subjects to participate. Written informed consent was obtained at the time of an in-person interview conducted in the subject's home. The study was approved by the Research Subjects Review Board of the University of Rochester.

Measures

Subjects were evaluated in their own homes by trained research interviewers. The Structured Clinical Interview for DSM-IV Psychiatric Disorders (SCID) (33) was used to determine the presence or absence of criteria sufficient to meet a diagnosis of current major depressive episode (MDE), and constituted the criterion standard for the analyses. Trained research personnel administered the SCID in the subject's home at the end of a comprehensive interview that lasted approximately 1.5 hours. The PHQ-9 and assessments of physical health and functional status, social supports, life circumstances, and service utilization preceded administration of the SCID. A final determination of MDE was made by consensus of the interviewers and a geriatric psychiatrist (YC). Given limited access to clients' medical

records or other objective measures of health status, we did not attempt to distinguish major affective disorder from MDE secondary to other causes (34). One hundred and one subjects (26.7%) met criteria for current MDE, 28 of whom were experiencing their first depressive episode; the remainder (n=73; 72.3%) had one or more prior episodes of major depression.

Psychometric analyses

We examined the sensitivity, specificity, positive (PPV) and negative predictive value (NPV), and the positive (LR+) and negative likelihood ratio (LR-) for each of three screening approaches. In this context, the sensitivity expresses the likelihood of screening positive with the PHQ if one has a MDE; the specificity indicates the likelihood of a negative screen if one does not have the illness. Reflecting the precision of the screening tool, the PPV indicates the proportion of subjects who screen positive that are correctly diagnosed by the SCID, while the NPV is the proportion of subjects with negative screening results who are correctly identified by the SCID as not having an MDE. Although useful as indicators of the probability that the screening tool accurately reflects the underlying condition, the PPV and NPV are dependent on the prevalence of MDE in the sample, and so may vary between sites. Therefore, we also include the LR+ and LR-, which are independent of prevalence rate. Expressed as a ratio, the LR+ (sensitivity/false positive rate) is the relative likelihood that a positive screen would be seen in someone *with* rather than *without* MDE; the LR- (false negative rate/specificity) is the relative likelihood that one would obtain a negative screen in someone *without* rather than *with* the illness.

The first screening strategy we employed was the PHQ-9 total score ("PHQ-9"; the overall score of all nine PHQ items; range = 0–27). Although the PHQ-9 may also be scored using a diagnostic algorithm to yield a categorical diagnostic rating, we chose the total score because the simplicity of its calculation and interpretation is better suited to a non-mental health, social service setting.

Second we examined the PHQ-2 total score ("PHQ-2", constituted simply by the first two questions of the PHQ-9; range = 0–6). We reasoned that if the shorter measure performed adequately with ASN clients, savings in administration time would benefit the agency and support uptake and dissemination. However, if the briefer tool were found to be too sensitive and insufficiently specific, then the high false positive rate would lead to unnecessary, time intensive, and costly referrals for evaluation and treatment.

Third, we tested a two-step process (referred to here as the "PHQ-2/9") in which response to the full nine-item version was calculated among only those subjects who screened positive to the PHQ-2 at a cut point of ≥ 2 . We chose a cut point of ≥ 2 for the first stage of the PHQ-2/9 approach in order to capture a larger number of cases than a higher initial cut point would have yielded (high sensitivity), hypothesizing that the second stage administration of the PHQ-9 at a cut point designed to yield greater specificity would strike an optimal balance. Note that ROC analysis can still be applied since the test results under this two-stage approach are naturally ordered. All PHQ-2 negatives (PHQ-2 < 2) constituted the lowest level above which all PHQ-2 positives are ordered according to their PHQ-9 scores.

Finally, we calculated the area under the receiver operating curve (AUC), which is an overall summary index of the discriminate ability, for each approach, and used cut-points designed to provide optimal sensitivity and specificity.

RESULTS

The psychometric properties for each screening approach are listed in Table 1. Figure 1 depicts their ROC curves. Optimal sensitivity and specificity for the PHQ-2 were 0.80 and

0.78 achieved at the standard cut point of 3. The AUC for the PHQ-2 was 0.87 (Table 2). For the PHQ-9, the optimal cut point appeared, like the PHQ-2, to correspond to the value generally recommended for primary care settings (22,23), a total score of ≥ 10 , yielding sensitivity and specificity of 0.82 and 0.87 respectively. As anticipated, the PHQ-9 resulted in comparable sensitivity but greater specificity than the PHQ-2, greater PPV, and greater LR+. The PHQ-9 had a higher overall AUC of 0.91.

Using a second stage (PHQ-9) cut point of ≥ 10 for the PHQ-2/9 approach (both PHQ-2 ≥ 2 and PHQ-9 ≥ 10), the sensitivity, specificity, PPV, NPV, LR-, and AUC were almost identical to those of the PHQ-9. The LR+ was marginally higher, indicating equal or better overall performance independent of the prevalence of MDE in the ASN agency's clientele. The large AUCs indicate that the three screening tools are useful in discriminating major depression (35). There is a statistically significant difference among the three AUCs (Chi-square: 10.66, DF: 2, $p=0.0049$), indicating that the PHQ-9 and PHQ-2/9 have better performance than the PHQ-2.

CONCLUSIONS

The findings of this study have significance for geriatric psychiatry because older adults resist seeking mental health care. Rather, strategies for detecting and treating mental illness in elders are best mounted in other sites in which patients are likely to present in high numbers, including primary care and ASN agencies. Collaborative depression care management models that integrate primary and mental health care have demonstrated effectiveness in improving outcomes of depressed seniors (36,37). Because social issues and stressful life events are predominant risk factors for late life depression, and because aging services providers are uniquely qualified to deal with these issues, the ASN is a natural partner with primary and mental health specialty care providers in delivery of care to depressed older adults. While the presence of mental illness among clients of both medical and non-medical home care agency clients is increasingly recognized as a public health problem, with rare exception (38) aging services agencies and providers of non-medical care management services have not been included in the design of collaborative systems of care for late life depression. Institution of routine screening for depression in the ASN setting is an important step towards further reducing barriers to mental health service delivery experienced by older adults, and to the integration of health and human services in providing their care.

The 26.7% prevalence of major depression in this sample of ASN care management clients is higher than some studies of community-dwelling older adults receiving home care services (17,18) and comparable to others (16). That over one quarter of clients had affective illness reinforces the need for tools with which agency providers can recognize and intervene to assure they receive mental health care. The PHQ-2 and the PHQ-9 both demonstrated good psychometric characteristics for detecting a current major depressive episode in this setting. The longer instrument may be relatively better suited to the ASN agencies though because its greater specificity will result in fewer false positives (non-depressed clients identified as cases) and thus reduce the cost to the agency and its elderly clients of unnecessary interventions.

The findings, however, further indicate that a two-stage process in which only those clients who screen positive on the PHQ-2 (scoring 2 or more) are administered the PHQ-9 performed at least as well, and perhaps better than either alternative; and the PHQ-2/9 would achieve those results with fewer questions than uniform screening with the PHQ-9, saving time and reducing respondent burden. More specifically, 165 of 378 subjects in this sample (43.6%) would have been spared the additional seven PHQ questions by using the PHQ-2/9

approach rather than administering the PHQ-9 to all. Given the average administration time of approximately five minutes for the PHQ-9 in this setting, reducing the screen to two questions (administered typically in a minute or less) would amount to meaningful time savings for the provider and reduce respondent fatigue in the course of a long and comprehensive care management assessment visit. In settings where the prevalence rate of MDE among clients is lower than we observed, of course, the proportion that would be spared the added questions would be higher and the savings in time and effort greater. For busy social service agencies operating on tight budgets while being responsible for addressing multiple, complex social needs of their elderly clients, such efficiencies may tip the balance in favor of adopting a depression screening program.

The tradeoff between using a tool and scoring method leading to a high level of sensitivity versus increased specificity is a decision that each agency must make based on its own unique circumstances. This decision would depend on the availability of staff comfortable with the administration and interpretation of the measure, patients' willingness to be referred, and access to primary care and/or mental health providers. Because screening questionnaires alone have little impact on case detection and treatment outcomes for depression (39), agencies must have mechanisms in place to assure that further indicated assessment and treatment are available to their depressed clients.

There are limitations that should be considered in interpreting these data. First, we have focused solely on use of the PHQ-9 for detection of major depressive illness in ASN clients. However, because significant stressors are a criterion for entry into the service system, milder or "subsyndromal" depression (SSD) is common in this population also. Most would agree that further evaluation and treatment of those who screen positive for major depression is clinically appropriate, and because SSD has been associated with significant functional morbidity (40), many would argue that it too should be a target for intervention. However, doing so would greatly increase demands on referral mechanisms; there is a lack of consensus regarding treatment guidelines for SSD; and it remains an empirical question whether the care managers' social interventions alone are sufficient to resolve the distress of less severely depressed elderly clients. Pending further study of the natural history of subsyndromal depression in clients receiving ASN interventions, we believe that screening and referral should target those with major depression.

Second, because the PHQ-9 was administered during the same interview and by the same interviewer as the diagnostic criterion standard SCID, there is the possibility of interviewer bias. However, it was our interpretation throughout the study that subjects tended to become more comfortable with the interviewer as the assessment progressed, and so tended to under-report their symptoms on the PHQ-9, which was administered early in the interaction. If there were bias associated with this aspect of the study design, it would tend to underestimate the psychometric properties of the screening tool. Also, the three screening approaches tested here were all derived from administration of the PHQ-9 rather than by administering the PHQ-2, PHQ-9, and their combination in unique samples of care management clients. For this reason we could not directly test the time saved by using the PHQ-2/9 instead of routine application of the full PHQ-9 scale. The measures may perform differently when administered by aging service providers than by research personnel, and results found here may not necessarily generalize to other aging services or older adult samples.

This study is the first to examine the criterion validity of the PHQ-2, the PHQ-9, and their sequential administration in aging services clients receiving in-home social work assessments. By developing stronger relationships with primary care and mental health systems, the ASN could help reduce the significant burden of late life mental disorders, and

in so doing would also help the ASN achieve its mission of maintaining independence and a higher quality of life for seniors. Use of the PHQ-2/9 in this setting can provide a means of identifying clients with clinically significant depressive disorders, and routing them to care. Future research is needed on the performance of the PHQ-2/9 when administered by ASN staff in the course of routine care management delivery; specific client, social service provider, and agency characteristics that may influence the performance, uptake, and implementation of these instruments; means by which to increase the sensitivity and specificity of depression screening tools, for example through use of more complex scoring algorithms that incorporate other client characteristics; and the best approaches to linking screening to effective treatment of affective illness in ASN clients.

Acknowledgments

The authors thank the entire Eldersource staff for their contributions and for making this work possible, and Constance Bowen and Judy Woodhams for their invaluable assistance in data collection.

Funding: Agency for Healthcare Research and Quality [T32HS000044 to T.R.]; National Institute of Mental Health [R24MH071604 to Y.C.]; and the American Foundation for Suicide Prevention [to Y.C.].

REFERENCES

1. Guilbert JJ. The world health report 2002 - reducing risks, promoting healthy life. *Education for Health* 2003;16:230. [PubMed: 14741909]
2. Greenberg PE, Kessler RC, Birnbaum HG, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry* 2003;64:1465–1475. [PubMed: 14728109]
3. Blazer DG. Depression in late life: review and commentary. *J Gerontol A Biol Sci Med Sci* 2003;58:249–265. [PubMed: 12634292]
4. Lyness JM, Caine ED, King DA, et al. Depressive disorders and symptoms in older primary care patients: one-year outcomes. *Am J Geriatr Psychiatry* 2002;10:275–282. [PubMed: 11994214]
5. Koenig HG, Meador KG, Cohen HJ, et al. Depression in elderly hospitalized patients with medical illness. *Arch Intern Med* 1988;148:1929–1936. [PubMed: 3415405]
6. Parmelee PA, Katz IR, Lawton MP. Depression among institutionalized aged: assessment and prevalence estimation. *J Gerontol* 1989;44:M22–M29. [PubMed: 2783434]
7. Mulsant BH, Ganguli M, Seaberg EC. The relationship between self-rated health and depressive symptoms in an epidemiological sample of community-dwelling older adults. *J Am Geriatr Soc* 1997;45:954–958. [PubMed: 9256848]
8. Conwell Y, Thompson C. Suicidal behavior in elders. *Psychiatr Clin North Am* 2008;31:333–356. [PubMed: 18439452]
9. Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. *J Affect Disord* 2002;72:227–236. [PubMed: 12450639]
10. Lyness JM, Heo M, Datto CJ, et al. Outcomes of minor and subsyndromal depression among elderly patients in primary care settings. *Ann Intern Med* 2006;144:496–504. [PubMed: 16585663]
11. Smit F, Ederveen A, Cuijpers P, et al. Opportunities for cost-effective prevention of late-life depression: an epidemiological approach. *Arch Gen Psychiatry* 2006;63:290–296. [PubMed: 16520434]
12. Nutting PA, Rost K, Dickinson M, et al. Barriers to initiating depression treatment in primary care practice. *J Gen Intern Med* 2002;17(2):103–111. [PubMed: 11841525]
13. Van Voorhees BW, Cooper LA, Rost KM, et al. Primary care patients with depression are less accepting of treatment than those seen by mental health specialists. *J Gen Intern Med* 2003;18:991–1000. [PubMed: 14687257]
14. Force USPST. Screening for depression: recommendations and rationale. *Ann Intern Med* 2002;136:760–764. [PubMed: 12020145]

15. O'Shaughnessy, CV. *The Aging Services Network: Accomplishments and Challenges in Serving a Growing Elderly Population*. Washington, DC: George Washington University; 2008.
16. Charlson ME, Peterson JC, Syat BL, et al. Outcomes of community-based social service interventions in homebound elders. *Int J Geriatr Psychiatry* 2008;23:427–432. [PubMed: 17918183]
17. Sirey JA, Bruce ML, Carpenter M, et al. Depressive symptoms and suicidal ideation among older adults receiving home delivered meals. *Int J Geriatr Psychiatry* 2008;23:1306–1311. [PubMed: 18615448]
18. Gum AM, Petkus A, McDougal SJ, et al. Behavioral health needs and problem recognition by older adults receiving home-based aging services. *Int J Geriatr Psychiatry* 2009;24:400–408. [PubMed: 18836987]
19. Casado BL, Quijano LM, Stanley MA, et al. Healthy IDEAS: Implementation of a Depression Program Through Community-Based Case Management. *Gerontologist* 2008;48:828–838. [PubMed: 19139256]
20. Mulrow CD, Williams JW Jr, Gerety MB, et al. Case-finding instruments for depression in primary care settings. *Ann Intern Med* 1995;122:913–921. [PubMed: 7755226]
21. Sharp LK, Lipsky MS. Screening for depression across the lifespan: a review of measures for use in primary care settings. *Am Fam Physician* 2002;66:1001–1008. [PubMed: 12358212]
22. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606–613. [PubMed: 11556941]
23. Lowe B, Unutzer J, Callahan C, et al. Monitoring Depression Treatment Outcomes With the Patient Health Questionnaire-9. *Medical Care* 2004;42:1194–1201. [PubMed: 15550799]
24. Ell K, Unutzer J, Aranda M, et al. Routine PHQ-9 depression screening in home health care: depression, prevalence, clinical and treatment characteristics and screening implementation. *Home Health Care Serv Quart* 2005;24:1–19.
25. Lamers F, Jonkers CC, Bosma H, et al. Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. *J Clin Epidemiol* 2008;61:679–687. [PubMed: 18538262]
26. American Psychiatric A. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition - Text Revision*. Washington, DC: American Psychiatric Association; 2000.
27. Lowe B, Kroenke K, Grafe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res* 2005;58:163–171. [PubMed: 15820844]
28. Whooley MA, Avins AL, Miranda J, et al. Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med* 1997;12:439–445. [PubMed: 9229283]
29. Foster MA, Ragsdale K, Dunne B, et al. Detection and treatment of depression in a VA primary care clinic. *Psychiatr Serv* 1999;50:1494–1495. [PubMed: 10543862]
30. Li C, Friedman B, Conwell Y, et al. Validity of the Patient Health Questionnaire 2 (PHQ-2) in identifying major depression in older people. *J Am Geriatr Soc* 2007;55:596–602. [PubMed: 17397440]
31. Eldersource Care Management Services. 2009 [accessed 12-30-09]. <http://www.eldersource.org/>
32. Council on Accreditation. 2009 [accessed 12-30-09]. <http://www.coanet.org/front3/front.cfm?view=7>
33. First, MB.; Spitzer, RL.; Gibbon, M., et al. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders - Non-patient Edition*. Biometrics Research Department, New York State Psychiatric Institute; 2001.
34. Koenig HG, George LK, Peterson BL, et al. Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. *Am J Psychiatry* 1997;154:1376–1383. [PubMed: 9326819]
35. Pines, JM.; Worth, WE. *The Epidemiology and Statistics of Diagnostic Testing*. West Sussex, UK: Blackwell Publishing; 2009.
36. Bruce ML, Ten Have T, Reynolds CF III, et al. Reducing Suicidal Ideation and Depressive Symptoms in Depressed Older Primary Care Patients: A Randomized Controlled Trial. *JAMA* 2004;291:1081–1091. [PubMed: 14996777]

37. Unutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* 2002;288:2836–2845. [PubMed: 12472325]
38. Quijano LM, Stanley MA, Petersen NJ, et al. Healthy IDEAS: A depression intervention delivered by community-based case managers serving older adults. *J Applied Gerontology* 2007;26:139–156.
39. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ* 2008;178:997–1003. [PubMed: 18390942]
40. Lyness JM, Kim J, Tang W, et al. The Clinical Significance of Subsyndromal Depression in Older Primary Care Patients. *Am J Geriatr Psychiatry* 2007;15:214–223. [PubMed: 17213374]

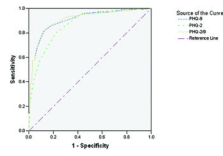


FIGURE 1.
ROC Curves for three approaches to depression screening

TABLE 1

Psychometric characteristics of three screening approaches

Screening Test	Cut Point*	Sensitivity	Specificity	PPV	NPV	LR+	LR-
PHQ-2	0.0	1.00	0.00	0.27		1.00	-
	1.0	0.99	0.26	0.33	0.99	1.34	0.04
	2.0	0.95	0.58	0.45	0.97	2.25	0.09
	3.0	0.80	0.78	0.57	0.91	3.63	0.26
	4.0	0.57	0.91	0.71	0.85	6.58	0.47
	5.0	0.40	0.96	0.80	0.82	11.08	0.62
	6.0	0.24	0.99	0.89	0.78	22.16	0.77
PHQ-9	7.0	0.93	0.62	0.47	0.96	2.45	0.11
	8.0	0.90	0.70	0.52	0.95	3.00	0.14
	9.0	0.86	0.82	0.63	0.94	4.67	0.17
	10.0	0.82	0.87	0.70	0.93	6.49	0.21
	11.0	0.74	0.91	0.75	0.91	8.20	0.29
	12.0	0.66	0.94	0.80	0.88	10.75	0.36
PHQ-2/9** (PHQ-2 ≥ 2, followed by PHQ-9)	7.0	0.92	0.71	0.54	0.96	3.18	0.11
	8.0	0.89	0.76	0.57	0.95	3.68	0.15
	9.0	0.85	0.84	0.66	0.94	5.25	0.18
	10.0	0.81	0.89	0.73	0.93	7.23	0.21
	11.0	0.73	0.92	0.76	0.90	8.80	0.29
	12.0	0.66	0.94	0.81	0.88	11.38	0.36

* Standard clinical cut point is bolded.

** Two-stage screen in which subjects who screen positive on PHQ-2 with total score ≥ 2 are then administered the full PHQ-9. Cut points apply to the second stage PHQ-9.

TABLE 2

Receiver Operating Characteristics (ROC)

Screening Test	Cut Point	Area Under Curve	Standard Error	Asymptotic Significance ^{***}	95% Confidence Interval	
					Lower Bound	Upper Bound
PHQ-2	PHQ-2 ≥ 3	0.869	0.020	<0.0001	0.830	0.908
PHQ-9	PHQ-9 ≥ 10	0.906	0.018	<0.0001	0.870	0.942
PHQ-2/9* (PHQ-2 ≥ 2 , followed by PHQ-9)	PHQ-9 ≥ 10	0.908	0.019	<0.0001	0.872	0.945

* Two-stage screen in which subjects who screen positive on PHQ-2 with total score ≥ 2 are then administered the full PHQ-9. Cut points apply to the second stage PHQ-9.

** Based on nonparametric Wilcoxon test