Published in final edited form as:
Am J Geriatr Psychiatry. 2009 July ; 17(7): 556–564.

Practical Depression Screening in Residential Care/Assisted Living: Five Methods Compared to Gold Standard Diagnoses

Lea C. Watson, MD, MPH[1], **Sheryl Zimmerman, PhD**[2,3], **Lauren W. Cohen, MA**[3], and **Rosalie Dominik, DrPH**[4]

[1]Department of Psychiatry of The University of North Carolina at Chapel Hill School of Medicine

[2]the School of Social Work at the University of North Carolina at Chapel Hill

[3]Cecil G. Sheps Center for Health Services Research at the University of North Carolina at Chapel Hill

[4]the Division of General Medicine and Epidemiology at the University of North Carolina at Chapel Hill

Abstract

Objective—To test the accuracy of five practical depression screening strategies in older adults residing in residential care/assisted living (RC/AL).

Design—Cross-sectional screening study.

Setting—Four RC/AL communities in North Carolina.

Participants—112 residents aged ≥ 65 and 27 staff members involved in their care.

Measurements—Direct care staff was trained in and completed the Cornell Scale for Depression in Dementia, modified for use by long-term care staff (CSDD-M-LTCS). They additionally responded to a one-item question 'Do you believe the resident is often sad or depressed?,' and the Minimum Data Set Depression Rating Scale (DRS). Residents responded directly to the Geriatric Depression Scale (15-item version; GDS-15) and the Personal Health Questionnaire, 2-item version (PHQ-2). A geriatric psychiatrist performed gold standard diagnostic interviews using the Structured Clinical Interview for DSM-IV. Sensitivities and specificities were calculated for all instruments at pre-determined cutpoints.

Results—Gold standard diagnoses yielded 14% prevalence of major or minor depression. The CSDD-M-LTCS and one-item screen completed by caregivers failed to significantly discriminate depressed cases. The DRS yielded high specificity (0.85), but low sensitivity (0.47). For the two resident reported measures, the PHQ-2 had a sensitivity of 0.80 and specificity of 0.71, and the GDS-15, 0.60 and 0.75 respectively.

Corresponding Author: Lea C. Watson MD, MPH, Department of Psychiatry, The University of North Carolina at Chapel Hill School of Medicine, CB 7160, Chapel Hill, NC 27599-7160, ph: 919-966-4390, fax: 919-966-2220, lea_watson@med.unc.edu.

The author contributions to the paper were as follows: Dr. Watson was involved in conception and design, acquisition of the data, project oversight, analysis and interpretation of data, drafting of the manuscript, and obtaining funding. Dr. Zimmerman was involved in conception and design, interpretation of the data, drafting of the manuscript, and obtaining funding. Ms. Cohen was involved in conception and design, acquisition of data, project oversight, data management, and drafting of the manuscript. Dr. Dominik was involved in the analysis and interpretation of data and drafting of the manuscript. All authors have given final approval to the submitted manuscript.

No authors report a conflict of interest.
No disclosures to report.

Conclusion—Measures completed by caregivers failed to adequately detect depression. Of the measures completed directly by residents, the PHQ-2 appears to have the best mix of brevity, sensitivity, and ease of administration.

Keywords

depression; screening; residential care/assisted living; long-term care; mental health

Residential care/assisted living (RC/AL) is a setting that has seen rapid growth in the United States as an alternative to nursing homes.1 These communities provide room, board, oversight, and assistance with activities of daily living, and although not required, some employ registered or licensed nursing staff.2 As RC/AL has grown, it has become evident that these settings provide care for impaired adults.3 In addition to the medical and functional impairment of these residents, the prevalence of clinically significant depression is notable and ranges from 13–25%.4,5 Unfortunately, studies in this population suggest that depression in RC/AL is largely unrecognized and associated with poor outcomes, including transfer to nursing home settings.6,7 Despite such cause for concern, RC/AL settings have no standard mechanism to screen for depression and limited personnel with training to conduct such screening. This deficiency merits attention because there are efficacious, feasible and low-risk treatments for depression in late-life.8,9 The challenge is to identify accurate and easily employed strategies to screen residents for depression, and then to implement service delivery models that work in this setting. The first step in this effort is to identify a practical screening strategy that effectively identifies residents with depression in RC/AL communities.

Multiple valid and reliable measures exist to screen for depression in older adults. In the RC/AL setting, where more than 50% of residents have cognitive impairment,10,11 such a measure must work despite the variation in cognitive ability that is inherent in this setting. One measure, the Cornell Scale for Depression in Dementia (CSDD) was specifically designed to assess depression in a cognitively impaired population. However, the CSDD was developed for use by clinicians who use caregiver and patient data to inform a unified score.12 It has not been tested for use by non-clinicians.

Because it is desirable to identify a screening tool that can be used by the paraprofessional staff who provide the majority of care in RC/AL, the intent of this study was to: (a) modify the CSDD in a way that makes it more easily used by direct care staff and assess the accuracy of the modified CSDD used by long-term care staff (referred to as the CSDD-M-LTCS) to detect depression against a gold standard; and (b) assess the accuracy of four other commonly used screening tools for depression, also as compared to a gold standard. The overall goal of the study was to identify effective screening strategies that are practical for use in RC/AL settings.

Methods

Sample

Data were collected from 112 residents and 27 direct care staff from 4 RC/AL settings participating in the Collaborative Studies of Long-term Care (CS-LTC). Facility eligibility was restricted to settings within 60 miles of the University of North Carolina that had 50 or more beds, did not primarily serve individuals with mental illness, and were not currently participating in another CS-LTC project. Ten eligible sites were identified and then invited to participate in random order until the desired sample size was reached. Five were contacted; one declined to participate, and the remaining 4 were enrolled in the project. Resident eligibility was limited to those who were aged 65 and older; did not have a

diagnosis of mental retardation or developmental disability; were English-speaking; and were considered healthy enough to participate for the duration of the project (approximately one month). A total of 151 residents were identified, and of these, 5 (3%) were ineligible. Of the remaining eligible residents, 26 (18%) declined participation and 8 (5%) could not be enrolled due to difficulty contacting the resident's family member for consent (which was done in the case of residents who had dementia and were unable to provide their own consent). The remaining 112 enrolled in the project, but 5 withdrew from the project prior to its conclusion. We report baseline (n=112) and complete (n= 107) data where indicated.

After a resident (or family) provided consent for participation, the facility administrator identified the staff member who was most involved in the care of that resident. That staff member was then invited to participate in the project and complete the CSDD-M-LTCS and provide other data in reference to the participating resident. In addition to the most involved, or “primary” staff member providing information, a second staff member who was also involved in the care of the resident was invited to participate for purposes of obtaining reliability data. Additional data were collected through interviews with staff and by chart review. Also, a psychiatrist conducted a clinical evaluation of the resident. All data were collected between November 2006 and June 2007. Study methods were approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Measures

The Cornell Scale for Depression in Dementia – Modified for Use by Long-Term Care Staff (CSDD-M-LTCS)—The original CSDD is a 19-item measure that uses patient and caregiver information to screen for depression in patients with dementia.12 The CSDD is intended for administration by a clinician, and assesses mood, behavior, physical signs, diurnal patterns, and ideational disturbances associated with depression over the past seven days. Each item of the CSDD is evaluated on a three-point scale (0-2) that reflects the severity of the symptom, for a total possible score of 38, with higher scores reflecting greater depression. The CSDD is currently the most accepted measure of depression in populations where dementia is prevalent, and has shown good sensitivity (90%) and specificity (75%) at a cutpoint of 12 for detecting major depression when administered by a clinician.12,13 In addition to dementia-specific populations, the clinician-administered CSDD has also been validated in populations without dementia14 and is commonly used in research settings.4,15

In order to make the CSDD accessible to non-clinicians, the form was modified based on cognitive testing with staff from a local nursing home, to remove technical language and to provide everyday examples of items. Further, the response options were modified from severity (i.e., none, mild, severe) to frequency (i.e., never or almost never, sometimes, most of the time). The 19 CSDD items were shown on the front of the form, and a complete explanation of each item was presented on the back. A scoring algorithm with step-by-step instructions for completion was also presented on the form. Sample items were all in reference to the stem “Over the last week, how often have you noticed the following …” and included sadness (sad face, sad voice, crying), slow behavior (moves slowly, talks slowly, slow to answer), and trouble falling asleep. The CSDD-M-LTCS is available upon request.

Staff training was conducted in either individual or small group sessions over 30 minutes. The training presentation and manual consisted of an overview of the CSDD-M-LTCS, an item-by-item guide to answering each question, and instruction in scoring the instrument. Staff also completed a practice CSDD-M-LTCS using a resident who recently left the community as an example.

To assess inter-rater reliability of the CSDD-M-LTCS, the primary and secondary staff caregiver completed the form for each resident. They were instructed to complete the form independently, although the instructions encouraged them to obtain whatever information would help them to accurately assess the resident's status (e.g., review the resident's chart, talk with staff on other shifts). Paired data were available for 111 residents. To assess test-retest reliability of the CSDD-M-LTCS, the primary staff member for each of the first 5 residents in the first three sites, and for the first 10 residents in the fourth site, was asked to complete the CSDD-M-LTCS twice, between two and five days apart. Paired test-retest data were available for 25 residents.

Other Measures Completed by Caregivers

One Item Screen: The primary staff caregiver for each resident was asked to respond to the Yes/No question *'Do you believe the resident is often sad or depressed?'*.16 Inter-rater and test-retest reliability was assessed using the same approach as that employed for the CSDD-M-LTCS.

Depression Rating Scale: The primary staff caregiver also completed the Minimum Data Set Depression Rating Scale (DRS)17 on behalf of each enrolled resident. This is a 7 item instrument designed for the nursing home environment that draws on continuous observation by staff. The caregiver was asked to respond based on the frequency of observing depressed behaviors over the last 30 days (0=not at all; 1=1-5 days per week; 2=daily or almost daily). A cutoff score of ≥3 on the DRS has been reported to yield 91% sensitivity and 69% specificity when tested against diagnoses of depression in a nursing home population.17

Measures Completed by Residents: A member of the research staff administered the 15-item version of the Geriatric Depression Scale (GDS-15)18 and the two-item Patient Health Questionnaire (PHQ-2)19 to enrolled residents, asking them to respond based on the past two weeks. In other studies of the GDS, sensitivity and specificity ranged from 79%-100% and 67-80% respectively in primary care elderly.20 The PHQ-2 (which asks residents whether they ever felt down, depressed or hopeless; or felt little interest or pleasure in doing things) has reported sensitivity of 100% and specificity of 77% and is considered appropriate for use in an older adult population.21 We scored the measure as positive if a resident answered 'yes' to having a depressed mood or anhedonia, regardless of the severity. This method is similar to the valid and reliable 2 item PRIME MD,22 but maintains a two week frame of reference (vs. one month) more in keeping with standard diagnostic procedure.

Other Measures: The research staff reviewed each enrolled resident's chart and obtained demographic, medical history, medical diagnoses, and physician contact information. The research staff also administered the Mini-Mental State Exam (MMSE)23 and collected data about residents' functional abilities using the Activities of Daily Living subscale of the Minimum Data Set (MDS-ADL).24 Staff members (both primary and secondary) also provided information about their own job title, training, length of employment, and demographics.

Psychiatric Evaluation—A geriatric psychiatrist (LW), blinded to all screening results, reviewed residents' charts and interviewed residents using the Structured Clinical Interview for DSM-IV.25 For all residents, the psychiatrist provided a diagnosis (or not) of current major or minor depression, and also whether or not her opinion was that the resident would be offered treatment for depression in a standard diagnostic circumstance regardless of diagnostic criteria.

Data Analysis—Prior to conducting the study, a sample size calculation determined that between 90 and 125 participants would provide reasonable precision for estimates of the likelihood ratio positive (LR+), assuming 15% prevalence of depression and using previously published estimates of the sensitivity and specificity for the CSDD and GDS-15.20

To assess the diagnostic accuracy of the five screening tools for predicting any (major or minor) depression, calculations were performed for the sensitivity, specificity and LR+ of each measure against the rendered diagnosis using pre-determined cut-points: CSDD-M-LTCS ≥ 8 and CSDD-M-LTCS ≥ 13; one item screen =1; DRS ≥3; GDS-15 ≥ 5; and PHQ-2 ≥ 1. Exact 95% confidence intervals (CI's) were determined for sensitivity and specificity and asymptotic 95% CI's for the LR+. For the CSDD-M-LTCS, DRS, and GDS-15 the area under the receiver operating curve (ROC) along with 95% confidence intervals based on the transformed AUCs were calculated.26 LR+ confidence intervals excluding 1, and also area under the curve (AUC) confidence intervals excluding 0.5, provide evidence that a screening tool has a statistically significant chance of discriminating between individuals with and without depression. Identical methods were used to assess the diagnostic accuracy of each of the five screening tools based on the overall impression of the psychiatrist as "would treat." The analyses described above were repeated after excluding the 11 participants with MMSE scores ≤ 10 in order to explore how inclusion of individuals with severe dementia may have affected our primary findings. Finally, intraclass correlation coefficients were used to assess reliability for the CSDD-M-LTCS and kappa coefficients to assess reliability of the one item screen. Reliability data were not obtained for the other measures.

Results

Sample Characteristics

Three of the RC/AL communities were for profit, and two (both for-profit) had a registered or licensed nurse on staff. The mean RC/AL community bedsize was 84 (SD 25); the average occupancy rate was 84% (SD 15%); the average monthly rate was $2762 (SD $1022).

Most residents were female (72%) and the average age was 83 (range 65 to 100) (Table 1). There was a range of cognitive and functional abilities, with an average MMSE of 20 and average impairment in 1.2 activities of daily living. A majority (70%) reported their health as good or better. Direct care staff (n=27) were all nursing assistants, nearly all female (96%), and predominantly African-American (85%).

The prevalence of depression based on all measures is shown in Table 2. Approximately one-quarter to one-third of residents screened positive for depression on the caregiver and resident reported measures (e.g., 36% on the CSDD-M-LTCS and 20% on the DRS). Clinician diagnoses yielded a 14% prevalence of major or minor depression, and increased to 22% when broadly considered as "would treat" despite the absence of meeting strict DSM-IV criteria.

Comparison of Screening Methods against Gold Standard

The performance of the CSDD-M-LTCS and the four other depression measures as a screening instrument are reported in Table 3. The gold standard comparator was a DSM-IV diagnosis of major or minor depression rendered by the psychiatrist. Extending the gold standard to include those whom the clinician "would treat" did not change the outcomes in any clinically meaningful way, nor did eliminating residents with MMSE ≤10 (more severe dementia).

CSDD-M-LTCS—The AUC for this measure was 0.66 (0.49,0.78) (Figure1), which did not reach statistical significance. Similarly, the likelihood ratios were not statistically significant at either cutpoint, with an LR+ of 1.3 (0.73-2.47) at a cutpoint of ≥8 and 2.4 (0.98, 0.5.66) at a cutpoint of ≥ 13. The sensitivity and specificity respectively were 0.47 (0.21,0.73) and 0.65 (0.55, 0.75) at a cutpoint of ≥8 ; 0.33 (0.12, 0.62) and 0.86 (0.77,0.92) at a cutpoint of ≥ 13. The test-retest reliability correlation was 0.83 (0.65,0.92), with an inter-rater correlation of 0.20 (0.01,0.37).

Other Caregiver Reported Measures—The one-item screen lacked significance for accuracy (based on the LR+ in Table 4), and the sensitivity and specificity were 0.47 (0.21, 0.73) and 0.74 (0.64,0.83), respectively. Test-retest and inter-rater reliability correlations were 0.70 (0.32,0.99) and 0.14 (-0.05, 0.33). The DRS had a statistically significant AUC of 0.70 (0.52,0.82) (Figure 1), with a sensitivity of 0.47 (0.21,0.73) and specificity of 0.85 (0.76,0.91) at the cutpoint of ≥3. Table 4 shows the sensitivity and specificity at additional selected cutpoints.

Measures Reported by Residents—The GDS-15 had an AUC of 0.80 (0.66,0.88), (Figure 1) reflecting its ability to accurately discriminate cases. It had a sensitivity of 0.60 (0.32,0.84) and a specificity of 0.75 (0.64,0.83) at the cutpoint of ≥5. Table 4 shows the sensitivity and specificity at additional selected cutpoints. The PHQ-2 had the best combination of sensitivity and specificity, 0.80 (0.52, 0.96) and 0.71(0.60,0.80), respectively, and showed statistically significant discriminating ability based on the LR+.

Discussion

The goal of our study was to test the accuracy of practical depression screening methods in the RC/AL setting, including the evaluation of the CSDD modified for use by direct care staff. We determined that three of five methods worked better than chance (95% CI of AUC did not include 0.5, and/or LR+ did not include 1): PHQ-2, GDS-15, and DRS. At predefined cutpoints the CSDD-M-LTCS did not significantly discriminate depressed individuals from other residents; the one item question also failed to discriminate significantly. Despite a high specificity of 0.85, the DRS had a low sensitivity (0.47), limiting its usefulness as a detection tool. The two remaining measures were completed by the residents themselves. The GDS-15 yielded a favorable AUC value of 0.80 (Figure 1), and had acceptable sensitivity (0.60) and specificity (0.75) at the traditional cutpoint of ≥ 5. Based on ROC data (Table 4), a cutpoint of ≥ 3 would optimize sensitivity (0.87), selectively improving the ability to “rule out” disease, which most closely approximates the goal of increasing the detection of potential cases. Similar findings have been reported in a community sample of “old-old,” where lowering the traditional cutpoint improved performance of the 30-item GDS.27 The resident-reported PHQ-2 had the best combination of sensitivity (0.80) and specificity (0.71) if one or both answers were positive. The PHQ-2 was scored as positive if either question was answered affirmatively (as opposed to a ranking of frequency for each symptom as in its original description)19 consistent with our effort to simplify the screening process, and still performed well.

Contrary to our hope that simplifying the CSDD or asking caregivers a simple one-item question would allow them to detect depression as “observers,” this was not the case. Although both the CSDD-M-LTCS and one-item question had good test-retest reliability, both also exhibited poor inter-rater reliability. These findings for caregiver completed measures may reflect that caregivers are not in tune with residents' depressive behaviors, or that there was variation in how well the residents were known by different caregivers. The latter is a real-world concern given the traditionally high turnover rate and erratic scheduling of the direct care workforce in this setting, and could logically account for an inability to

recognize a pattern of behaviors.28 It is also possible that the scale itself is difficult to follow for caregivers without formal clinical training, as is common for nursing assistants. It was our own anecdotal experience trying to use the original CSDD with such staff – and their general resistance – that led us to simplify the wording and scoring for this study; however, it is possible that these modifications may have contributed to its poor performance. Another source of potential discordance, which was not examined in this project, was that between observers and the residents themselves, which has been reported in other work.29 In an RC/AL population, discordance might be expected due to the high prevalence of functional and physical disability,30 which makes a distinction between the "physical" and the "mental" more difficult.

Self-reported measures of depression inherently eliminate observer biases, reflecting the subjective experience of the resident. Many brief measures have been developed and tested in older adults in multiples settings,31-34 but not specifically in RC/AL. The decision to test the GDS-15 and PHQ-2 in this setting represents an effort to achieve the least common screening burden for vulnerable residents with instruments familiar to most U.S. providers. Their ease of use accommodates self-administration as well as verbal administration by care staff as necessary, due to vision or functional impairment. It is especially heartening to know that the two- item PHQ-2 performs best in this setting.

Limitations of this study include the relatively small sample size and event rate (n=107; prevalence of any depression 14%), which limited the precision of point estimates and our power to detect true discriminating ability of any measure. This is reflected in wide confidence intervals, and findings should be interpreted with this in mind. Also, all subjects were from one region of central North Carolina, although there has been no indication in the literature that the widely-used measures we evaluated are regionally or culturally biased. We also elected to evaluate residents who have dementia for depression using these same standard diagnostic procedures, despite the limitations of "gold standard" validity in cognitive impairment. This is an unresolved issue in the field that is receiving increased attention,35,36 but in the absence of clear guidelines, our intent was to be as inclusive as possible and mimic typical clinical practice. It is noteworthy that removing the 11 residents with severe dementia from the analyses did not change our results.

As many as 25% of RC/AL residents are depressed, but most go unrecognized. We found three methods of screening for depression to have statistically significant discriminating ability in this setting, – the PHQ-2, GDS-15 and DRS, but the PHQ-2 had the best sensitivity. The PHQ-2 is also the most practical because of its brevity and ease of use. Although slightly longer, the GDS-15 is easy to administer and could play a role in detection at a lower cutpoint, especially in RC/AL settings that may have already adopted its use.

Successful screening can improve detection of depression.37 Essential steps will need to be taken to translate screening findings into meaningful outcomes in the widely heterogeneous RC/AL setting, where access to psychiatric care is highly variable. Staff and administrators must "buy-in" and adopt routine depression screening, ideally by resident self-administration, at admission and with periodic surveillance; positive screens must prompt a well-defined referral and/or treatment response; and recommended treatments must be monitored and appropriately adjusted over time in line with the current goals of care. Future research directions following from this effort include testing a practical depression screening, referral, and treatment intervention in RC/AL that honors resident autonomy while also providing an opportunity to reduce the suffering caused by this common and debilitating disease.

Acknowledgments

The authors thank the staff, residents, and family members who participate in the Collaborative Studies of Long-Term Care (CS-LTC) for their time and commitment to learning more about and improving the quality of life in assisted living/residential care communities and nursing homes.

Financial support was provided by a Multidisciplinary Clinical Research Career Development Award to Lea C. Watson, MD, MPH (1K12HD052191) and a National Institute on Aging grant to Sheryl Zimmerman, PhD (K02 AG00970). Rosalie Dominik is supported by grant (UL1RR025747) from the Clinical and Translational Science Award of the Division of Research Resources, National Institutes of Health. The sponsors did not play a role in the preparation of this paper or the data presented herein.

References

1. Spillman BC, Liu K, McGilliard C. Trends in Residential Long-Term Care: Use of Nursing Homes and Assisted Living and Characteristics of Facilities and Residents. Nov.2002 Available at: http://aspe.hhs.gov/daltcp/reports/rltct.pdf
2. Stearns SC, Park J, Zimmerman S, et al. Determinants and effects of nurse staffing intensity and skill mix in residential care/assisted living settings. Gerontologist. 2007; 47:662–71. [PubMed: 17989408]
3. Golant SM. Do impaired older persons with health care needs occupy U.S. assisted living facilities? An analysis of six national studies. J Gerontol B Psychol Sci Soc Sci. 2004; 59:S68–79. [PubMed: 15014094]
4. Watson LC, Garrett JM, Sloane PD, et al. Depression in assisted living: results from a four-state study. Am J Geriatr Psychiatry. 2003; 11:534–42. [PubMed: 14506087]
5. Watson LC, Lehmann S, Mayer L, et al. Depression in assisted living is common and related to physical burden. Am J Geriatr Psychiatry. 2006; 14:876–83. [PubMed: 17001027]
6. Gruber-Baldini AL, Zimmerman S, Boustani M, et al. Characteristics associated with depression in long-term care residents with dementia. Gerontologist. 2005; 45(1):50–5. [PubMed: 16230749]
7. Zimmerman S, Sloane PD, Eckert JK, et al. How Good Is Assisted Living? Findings and Implications From an Outcomes Study. J Gerontol B Psychol Sci Soc Sci. 2005; 60:S195–204. [PubMed: 15980295]
8. Reynolds CF, Dew MA, Pollock BG, et al. Maintenance Treatment of Major Depression in Old Age. N Engl J Med. 2006; 354:1130–8. [PubMed: 16540613]
9. Shanmugham B, Karp J, Drayer R, et al. Evidence-based pharmacologic interventions for geriatric depression. Psychiatr Clin North Am. 2005; 28:821–35. viii. [PubMed: 16325731]
10. Rosenblatt A, Samus QM, Steele CD, et al. The Maryland Assisted Living Study: prevalence, recognition, and treatment of dementia and other psychiatric disorders in the assisted living population of central Maryland. J Am Geriatr Soc. 2004; 52:1618–25. [PubMed: 15450036]
11. Zimmerman S, Sloane PD, Williams CS, et al. Residential care/assisted living staff may detect undiagnosed dementia using the minimum data set cognition scale. J Am Geriatr Soc. 2007; 55:1349–55. [PubMed: 17767676]
12. Alexopoulos G, Abrams R, Young R, et al. Cornell Scale for Depression in Dementia. Biological Psychiatry. 1988; 23:271–84. [PubMed: 3337862]
13. Vida S, Des Rosiers P, Carrier L, et al. Depression in Alzheimer's Disease: Receiver Operating Characteristic Analysis of the Cornell Scale for Depression in Dementia and the Hamilton Depression Scale. J Geriatr Psychiatry Neur. 1994; 7:159–62.
14. Alexopoulos G, Abrams R, Young R, et al. Use of the Cornell Scale in Nondemented Patients. J Am Geriatr Soc. 1988; 36:230–6. [PubMed: 3339232]
15. Kopetz S, Steele CD, Brandt J, et al. Characteristics and outcomes of dementia residents in an assisted living facility. Int J Geriatr Psychiatry. 2000; 15:586–93. [PubMed: 10918338]
16. Mahoney J, Drinka TJ, Abler R, et al. Screening for depression: single question versus GDS. J Am Geriatr Soc. 1994; 42:1006–8. [PubMed: 8064088]
17. Burrows AB, Morris JN, Simon SE, et al. Development of a minimum data set-based depression rating scale for use in nursing homes. Age Ageing. 2000; 29:165–72. [PubMed: 10791452]

18. Sheik, J.; Yesavage, J. Geriatric DepressionScale (GDS): recent findings and development of a shorter version. New York: Howarth Press; 1986.
19. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care. 2003; 41:1284–92. [PubMed: 14583691]
20. Watson L, Pignone M. Screening accuracy for late-life depression in primary care: a systematic review. J Fam Pract. 2003; 52:956–964. [PubMed: 14653982]
21. 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]
22. 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; 12:439–45. [PubMed: 9229283]
23. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12:189–98. [PubMed: 1202204]
24. Morris JN, Fries BE, Morris SA. Scaling ADLs within the MDS. J Gerontol A Biol Sci Med Sci. 1999; 54:M546–53. [PubMed: 10619316]
25. Apa. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders. 4th. Washington, D.C.: American Psychiatric Association; 1994.
26. Zhou, XH.; McClish, DK.; Obuchowski, NA. Statistical Methods in Diagnostic Medicine. Wiley; 2002.
27. Watson LC, Lewis CL, Kistler CE, et al. Can we trust depression screening instruments in healthy ‘old-old’ adults? Int J Geriatr Psychiatry. 2004; 19:278–85. [PubMed: 15027044]
28. Zimmerman, S.; Sloane, PD.; Eckert, JK. Assisted living : needs, practices, and policies in residential care for the elderly. Baltimore, MD: Johns Hopkins University Press; 2001.
29. Burrows AB, Satlin A, Salzman C, et al. Depression in a long-term care facility: Clinical features and discordance between nursing assessment and patient interviews. J Am Geriatr Soc. 1995; 43(10):1118–1122. [PubMed: 7560702]
30. Jang Y, Bergman E, Schonfeld L, et al. Depressive symptoms among older residents in assisted living facilities. Int J Aging Hum Dev. 2006; 63(4):299–315. [PubMed: 17191436]
31. Van Marwijk H, Arnold I, Bonnema J, et al. Self-report depression scales for elderly patients in primary care: A preliminary study. Fam Pract. 1993; 10(1):63–65. [PubMed: 8477896]
32. Weyerer S, Killman U, Ames D, et al. The Even Briefer Assessment Scale for Depression (EBAS DEP): Its suitability for the elderly in geriatric care in English- and German-speaking countries. Int J Geriatr Psychiatry. 1999; 14(6):473–480. [PubMed: 10398358]
33. Lim PP, Ng LL, Chiam PC, et al. Validation and comparison of three brief depression scales in an elderly Chinese population. Int J Geriatr Psychiatry. 2000; 15(9):824–830. [PubMed: 10984729]
34. D'Ath P, Katona P, Mullan E, et al. Screening, detection and management of depression in elderly primary care attenders I: The acceptability and performance of the 15-item Geriatric Depression Scale (GDS15) and the development of short versions. Fam Pract. 1994; 11(3):260–266. [PubMed: 7843514]
35. Lyketsos CG, DelCampo L, Steinberg M, et al. Treating depression in Alzheimer disease: efficacy and safety of sertraline therapy, and the benefits of depression reduction: the DIADS. Arch Gen Psychiatry. 2003; 60:737–46. [PubMed: 12860778]
36. Olin JT, Katz IR, Meyers BS, et al. Provisional diagnostic criteria for depression of Alzheimer disease: rationale and background. Am J Geriatr Psychiatry. 2002; 10:129–41. [PubMed: 11925274]
37. Pignone MP, Gaynes BN, Rushton JL, et al. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2002; 136(10):765–776. [PubMed: 12020146]

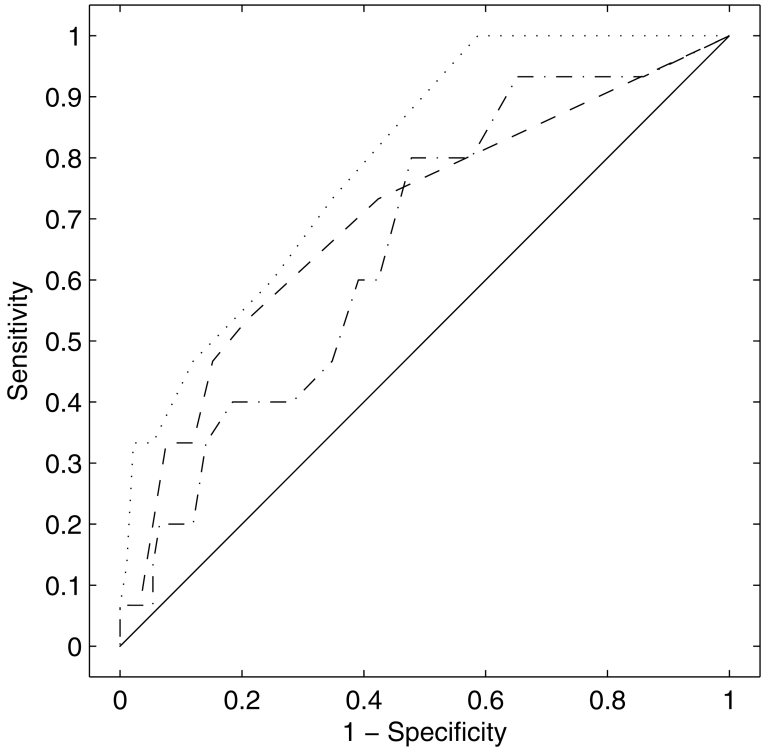

Figure 1.
ROC curves for CSDD-M-LTCS, DRS, and GDS-15 (Chance on Diagonal)
Dash-Dot Line (_ · _ · _ ·): The Cornell Scale for Depression in Dementia – Modified for Use by Long – Term Care Staff (CSDD-M-LTCS)
Dashed Line (_ _ _): Depression Rating Scale (DRS)
Dotted Line (· · · · · ·): Geriatric Depression Scale (GDS-15)

Table 1

Sample characteristics

Characteristic	N (%) or Mean (SD)
Residents (N=112)	
Female	81 (72)
Age	83 (8.1)
White	94 (84)
Minimum high school education	88 (79)
Any college	49 (44)
Years in facility	2.1 (2.2)
MMSE score	20.0 (7.1)
Number of ADL * dependencies	1.2 (1.93)
Self-rated health good or better	76 (70)
Staff (N=27)	
Female	26 (96)
Age	37 (12.4)
African-American	23 (85)
Years in present job	1.6 (1.96)

* ADL = Activity of daily living

Table 2

Prevalence of depression by different assessment methods (n=107)

Measure	Prevalence (%)
Caregiver reported	
CSDD-M-LTCS ≥ 8 *	36
One-item: often sad or depressed?	29
Depression Rating Scale ≥ 3 *	20
Resident reported	
GDS-15 ≥ 5 *	30
PHQ-2 ≥1	37
Clinician reported	
“Any” depression (major or minor)	14
“Would treat”	22

* The mean score of the CSDD-M-LTC was 6.7 (SD 6.2); of the Depression Rating Scale was 1.4 (SD 2.3); and of the GDS-15 was 3.1 (SD 2.7).

Table 3

Performance of screening measures compared to DSM-IV diagnoses of major or minor depression

Screening measure (n)	AUC * (95% CI)	Cut-point	Sensitivity (95% CI†)	Specificity (95% CI†)	LR± positive (95% CI†)
Caregiver completed measures					
CSDD-M-LTCS (107)	0.66 (0.49-0.78)	≥8	0.47 (0.21-0.73)	0.65 (0.55-0.75)	1.3 (0.73-2.47)
		≥13	0.33 (0.12-0.62)	0.86 (0.77-0.92)	2.4 (0.98-5.66)
One-item screen (107)	-	-	0.47 (0.21-0.73)	0.74 (0.64-0.83)	1.8 (0.94-3.40)
Depression Rating Scale (107)	0.70 (0.52-0.82)	≥3	0.47 (0.21-0.73)	0.85 (0.76-0.91)	3.1 (1.49-6.33)
Resident completed measures					
GDS-15 (107)	0.80 (0.66-0.88)	≥5	0.60 (0.32-0.84)	0.75 (0.65-0.83)	2.4 (1.39-4.14)
PHQ-2 (104)	-	≥1	0.80 (0.52-0.96)	0.71 (0.60-0.80)	2.7 (1.82-4.13)

*AUC = Area under the curve

†CI = Confidence interval

±LR = Likelihood ratio

Table 4

Selected cutpoints for two instruments based on Receiver Operator Characteristic (ROC) analyses

Cutpoint	Sensitivity	Specificity
Depression Rating Scale		
≥ 1	0.73	0.58
≥ 2	0.53	0.80
≥ 3	0.47	0.85
≥ 4	0.33	0.88
GDS-15		
≥ 3	0.87	0.53
≥ 4	0.73	0.65
≥ 5	0.60	0.75
≥ 6	0.47	0.88

Note: See Figure 1 for Receiver Operator Characteristics (ROC) curve.