

Application of a stratum-specific likelihood ratio analysis in a screen for depression among a community-dwelling population in Japan

Norio Sugawara^{1,2}Ayako Kaneda²Ippei Takahashi³Shigeyuki Nakaji³Norio Yasui-Furukori²

¹Department of Clinical Epidemiology, Translational Medical Center, National Center of Neurology and Psychiatry, Kodaira, Tokyo, ²Department of Neuropsychiatry, Hirosaki University School of Medicine, Hirosaki,

³Department of Social Medicine, Hirosaki University School of Medicine, Hirosaki, Japan

Background: Efficient screening for depression is important in community mental health. In this study, we applied a stratum-specific likelihood ratio (SSLR) analysis, which is independent of the prevalence of the target disease, to screen for depression among community-dwelling individuals.

Method: The Center for Epidemiologic Studies Depression Scale (CES-D) and the Mini International Neuropsychiatric Interview (MINI) were administered to 789 individuals (19–87 years of age) who participated in the Iwaki Health Promotion Project 2011. Major depressive disorder (MDD) was assessed using the MINI.

Results: For MDD, the SSLRs were 0.13 (95% CI 0.04–0.40), 3.68 (95% CI 1.37–9.89), and 24.77 (95% CI 14.97–40.98) for CES-D scores of 0–16, 17–20, and above 21, respectively.

Conclusion: The validity of the CES-D is confirmed, and SSLR analysis is recommended for its practical value for the detection of individuals with the risk of MDD in the Japanese community.

Keywords: screening, depression, Center for Epidemiologic Studies Depression Scale, stratum-specific likelihood ratio

Introduction

Major depression is a serious, recurrent mental disorder with high lifetime prevalence worldwide.¹ Major depressive disorder (MDD) constitutes a crucial public health burden. In Japan, MDD has a lifetime prevalence of 6.2% and a 12-month prevalence of 2.2%, is more prevalent among females than among males.² The World Health Organization has projected that MDD will become the number one leading cause of worldwide disability-adjusted life-years (DALYs) by 2020.³

More than 90% of suicide victims were retrospectively diagnosed with a psychiatric problem at their time of committing suicide, and approximately two-thirds of suicide victims were diagnosed with depression.^{4,5} Prior studies have identified depression as the major risk factor for suicidal ideation, suicide attempts, and successful suicides.^{6–8} Given the strong relationship between suicidal behavior and depression, screening for and treating depressive disorder have been proposed as one approach to prevent suicide.⁹

The Center for Epidemiologic Studies Depression Scale (CES-D) is a relatively simple and quick assessment that is inexpensive to administer and nonhazardous to patients; thus, this scale is acceptable for screening a large population.¹⁰ Shima et al¹¹ first reported the clinical validity of the CES-D Japanese translation and its

Correspondence: Norio Sugawara
Department of Clinical Epidemiology,
Translational Medical Center, National
Center of Neurology and Psychiatry,
Kodaira, Tokyo 187-8551, Japan
Tel +81 42 341 2712
Fax +81 42 346 3503
Email nsuga3@yahoo.co.jp

standard cutoff point of 15/16. Subsequently, other studies from Japan have supported this finding and shown a good level of criterion-related validity and internal reliability in a sample of Japanese workers.^{12,13} Although continuous CES-D scores are informative, interpreting screening results by utilizing score categories high predictive values with respect to the administration of efficient interventions is a convenient method.

Stratum-specific likelihood ratios (SSLRs) could be used to obtain such categories.¹⁴ The likelihood ratio (LR) provides a direct estimate of how much a test result will change the odds of having a disease and incorporates both the sensitivity and specificity of the test.¹⁵ SSLRs are calculated by dividing the continuous likelihood ratios into strata.¹⁶ However, reports regarding the use of SSLR analysis to screen for depressive disorders with the CES-D among community populations.^{17,18}

The objective of this study was to determine the CES-D score categories that have predictive clinical value for community-based screening for depression.

Method

Participants

The subjects were 789 volunteers (19–87 years of age, 289 males and 500 females) who participated in the Iwaki Health Promotion Project 2011.¹⁹ These individuals were residents of the Iwaki district, which is a rural area of the city of Hirosaki, in northern Japan. Iwaki is a stable community with a population of 11,863 individuals. The mean age of the participants was 57.8 years (with SD of 11.4 years and a range of 19–87 years). The data collection procedures for this study were approved by the ethics committee of the Hirosaki University School of Medicine, and all subjects provided written informed consent before participating in this project. Demographic data (age and gender) were obtained from self-reported questionnaires and interviews.

Procedure

The Japanese version of the CES-D was administered to all participants to measure depressive symptoms.^{10,20} The CES-D is a 20-item, self-reported measure that focuses on depressive symptoms during the week prior to the administration of the questionnaire. The maximum score on this scale is 60, and a CES-D score of 16 or more is regarded as indicative of the presence of depression.

The Mini International Neuropsychiatric Interview (MINI) is a short, structured diagnostic interview for

psychiatric disorders in the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV) and the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10).^{21,22} In this study, we applied the portion of the MINI that is used to identify a major depressive episode. Among participants who responded “Yes” to A1 (depressed mood) and/or A2 (loss of interest), MDD was defined as a score of ≥ 5 on this portion of the major depressive episode section of the MINI.

Statistical analysis

The sensitivity, specificity, positive predictive value, and negative predictive value were calculated for several cutoff scores, as well as for the traditional cutoff score of 16 on the CES-D scale. Receiver operating characteristic (ROC) analysis was performed, with the true-positive rate (sensitivity) plotted on the vertical axis and the false-positive rate ($1 - \text{specificity}$) plotted on the horizontal axis; this approach allows display of all pairs of sensitivity and specificity values achievable as the threshold is changed from a low score to a high score.²³ The area under the curve (AUC) can be used as a quantitative indicator of the information content of a test. An AUC of 1.0 indicates perfect accuracy, whereas an AUC of 0.5 indicates a nondiscriminating test. The software used for the ROC analysis was Statistical Package for the Social Sciences (SPSS), version 24. The ROC curve is also used to determine the score that maximizes a screening test's efficacy. The point on the ROC curve with the shortest distance to the intersection of the sensitivity and $1 - \text{specificity}$ values on the ROC graph is defined to be the optimal cutoff score.

We computed the SSLRs and 95% CIs using a program developed by Peirce and Cornell.²⁴ An SSLR indicates how much more likely or less likely a specific test result is for individuals with a disease than for individuals without this disease; thus, this ratio could reveal the efficiency of a screening test. LRs > 10 and < 0.1 indicate strong evidence for diagnosis and exclusion of diseases in clinical practice, respectively.¹⁴ To achieve the optimum number of strata, we followed the rules proposed by previous studies^{24,25} as follows: 1) provide sufficient disordered and nondisordered subjects in each stratum to allow the SSLRs to be monotonically related; and 2) collapse strata where the SSLRs are close to one another and their 95% CIs easily overlap.

Logistic regression models were used to test whether the strata provided significantly more information than

a single cutoff point.²⁶ First, a logistic regression model with a dichotomous predictor was fitted. Then, a second model with the same dichotomous predictor and the stratum as the categorical predictor was fitted. The difference between these two values was analyzed using the chi-square statistic under the null hypothesis that the strata predictor did not add more predictive ability than the single cutoff point.

Results

The overall scale was found to be reliable ($\alpha = 0.77$), and 12% (95/789) of the participants had a CES-D score of at least 16.

The sensitivities, specificities, positive predictive values, and negative predictive values are presented in Table 1. The ROC curve for the CES-D is depicted in Figure 1. The cutoff value, which was determined based on the shortest distance between any point on the ROC curve and the upper left intersection of the sensitivity and 1 – specificity values on the ROC graph, was 16. The AUC calculated using the ROC analysis was 0.98 (95% CI 0.96–1.00; $p < 0.001$).

The SSLR analysis results are indicated in Table 2. The recommended SSLRs determined for MDD were 0.13 (95%

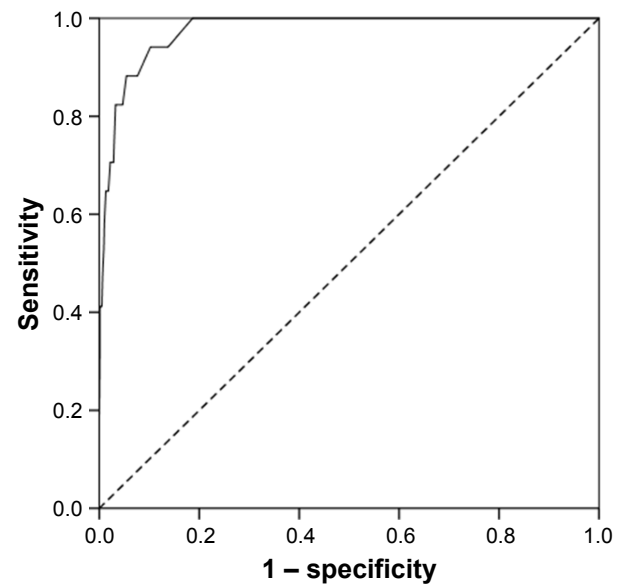


Figure 1 ROC curve for depressive disorder determined using the MINI and the CES-D.

Abbreviations: CES-D, Center for Epidemiologic Studies Depression Scale; MINI, Mini International Neuropsychiatric Interview; ROC, receiver operating characteristic.

CI 0.04–0.40), 3.68 (95% CI 1.37–9.89), and 24.77 (95% CI 14.97–40.98) for CES-D scores of 0–16, 17–20, and 21–60, respectively. Given the prior probability of the base rate of MDD used in the present study (2.1%), CES-D scores falling in the range of 0–16 had an SSLR < 1.0 and thus shifted the posttest probability of having that disorder to a very low level (0.29%).

For MDD, the $-2 \cdot \log$ LR for the model with a single cutoff point (cutoff of 15/16) was 101.2. The $-2 \cdot \log$ LR for the model with the stratum as a categorical predictor was 93.0. The difference between these two values was 8.3. The corresponding p -value was small ($p = 0.004$), indicating significant improvement when the stratum for MDD was used.

Discussion

Our results in which the CES-D had an AUC value of 0.98 demonstrated that this scale had high validity for the detection of MDD evaluated by the MINI in a community-dwelling population. Furthermore, a score of 16, which is traditionally

Table 1 Sensitivity, specificity, and positive and negative predictive values of the CES-D

Cutoff point	Sensitivity	Specificity	Positive predictive value	Negative predictive value
13/14	1.000	0.813	0.106 (17/161)	1.000 (628/628)
14/15	0.941	0.863	0.131 (16/122)	0.999 (666/667)
15/16	0.941	0.898	0.168 (16/95)	0.999 (663/694)
16/17	0.882	0.924	0.203 (15/74)	0.997 (713/715)
17/18	0.882	0.946	0.263 (15/57)	0.997 (730/732)
18/19	0.824	0.953	0.280 (14/50)	0.996 (736/739)
19/20	0.824	0.968	0.359 (14/39)	0.996 (747/750)
20/21	0.706	0.972	0.353 (12/34)	0.993 (750/755)
21/22	0.706	0.978	0.414 (12/29)	0.993 (755/760)
22/23	0.647	0.982	0.440 (11/25)	0.992 (758/764)
23/24	0.647	0.987	0.524 (11/21)	0.992 (762/768)
24/25	0.588	0.990	0.556 (10/18)	0.991 (764/771)
25/26	0.529	0.991	0.563 (9/16)	0.990 (765/773)
26/27	0.471	0.994	0.615 (8/13)	0.988 (767/776)
27/28	0.412	0.995	0.636 (7/11)	0.987 (768/778)
28/29	0.412	0.995	0.636 (7/11)	0.987 (768/778)
29/30	0.412	0.997	0.700 (7/10)	0.987 (769/779)
30/31	0.412	0.997	0.778 (7/9)	0.987 (770/780)
31/32	0.412	0.999	0.875 (7/8)	0.987 (771/781)
32/33	0.353	0.999	0.857 (6/7)	0.986 (771/782)
33/34	0.176	1.000	1.000 (3/3)	0.982 (772/786)

Abbreviation: CES-D, Center for Epidemiologic Studies Depression Scale.

Table 2 SSLRs obtained for MDD in this study

CES-D score	MDD, n	Non-MDD, n	SSLR	95% CI	Posttest probability, %
0–16	2	713	0.13	(0.04–0.40)	0.29
17–20	3	37	3.68	(1.37–9.89)	7.50
21–60	12	22	24.77	(14.97–40.98)	35.29

Abbreviations: CES-D, Center for Epidemiologic Studies Depression Scale; MDD, major depressive disorder; SSLR, stratum-specific likelihood ratio.

used as the cutoff score, exhibited the highest total sensitivity and specificity (0.941 and 0.898, respectively). However, the positive predictive value (0.168) for this cutoff, which depends on the prevalence of the examined disease, indicates that many of the positive results obtained using this testing procedure are false positives. Given the prevalence of depression in prior community-based studies,^{2,9} measures that are independent of the prevalence are needed for community-based screening.

In this study, we utilized the SSLR approach, which is independent of the target disease prevalence, and obtained three CES-D strata. By accounting for an individual's prior probability, this approach could provide a basis to calculate the posttest probability in each specific stratum. Our results demonstrated that higher CES-D scores were associated with higher risks of MDD. Individuals with a CES-D score of 21 or greater had an SSLR that was significantly >10. Relative to the other categories, these score categories will provide primary care physicians and psychiatrists with greater opportunities to detect individuals with MDD via further investigation following screening using the CES-D.

A prior study conducted in first-visit psychiatric patients aged 36.9 ± 16.0 (mean $\pm$ SD) years classified CES-D scores into three strata of <29, 30–49, and >50; the SSLRs for these three strata were 0.35 (95% CI 0.25–0.49), 2.3 (95% CI 1.8–3.1), and 11.7 (95% CI 3.1–44.0), respectively.²⁵ Another study conducted in Japan identified three strata <16, 17–19, and >20 for working individuals aged 42.0 ± 11.4 (mean $\pm$ SD) years; the SSLRs for these three strata were 0.06 (95% CI 0.02–0.18), 1.9 (95% CI 1.78–4.62), and 12.4 (95% CI 10.2–15.1), respectively.¹⁷ However, an imbalanced gender ratio in that study (males: $n=1,868$ and females: $n=351$) may have influenced the results. Among Taiwanese adolescents aged 12–16 years, Yang et al¹⁸ showed three strata of <28, 29–48, and >49. The SSLRs for these three strata were 0.63 (95% CI 0.42–0.92), 4.0 (95% CI 2.4–6.8), and 11.8 (95% CI 3.4–50.9), respectively. These studies differed in the mean age, gender ratio, and setting (clinical or community), which might have affected the depressive symptomatology. Furthermore, differences concerning the assessment of depressive disorder (DSM-III-R or Kiddie Schedule for Affective Disorders and Schizophrenia [K-SADS]) could affect the CES-D score categories. Although differences among the score categories have been identified to date, understanding the existence of a dose-dependent relationship between the CES-D score and the risk of MDD. Compared with previous studies, our score categories of the CES-D

might be useful for screening the older population in a community for depression.^{17,18,25}

We must consider the interpretation of the false-positive results obtained when screening using the CES-D. Prior studies have found poor positive predictive power associated with using the CES-D to screen for depression.^{27,28} False positives might be induced by active medical conditions, alcohol-related disorders, and/or dysthymia.²⁹ Although individuals with false-positive results do not satisfy the criteria for MDD, they could benefit from follow-up with primary care physicians or psychiatrists. We should also consider false-negative cases. Individuals who were afraid of being diagnosed with MDD or who were unaware of their illness might not express their symptoms during the screening process.³⁰ An objective biomarker for diagnosing MDD is needed to identify these individuals.

Although screening for depression is important, it is also necessary to provide consistent treatment and follow-up. Merely identifying individuals with depression would not be beneficial. Furthermore, recent evidence has raised questions regarding the degree to which standard treatments for depression benefit patients who are identified via screening.³¹ A meta-analysis reported that the benefits of antidepressant medication compared with those of a placebo were minimal or nonexistent in patients with mild or moderate depressive symptoms.^{32,33} Therefore, a combination of prompt detection and treatment appropriate for the symptom severity is needed for effective community-based screening for depression.

This study has certain limitations. First, all of the participants were volunteers who were interested in their health and may therefore have been healthier than the general population. Thus, the community members who were not included in the study may have experienced different depressive symptoms than the study participants. This type of “selection bias” must be considered in studies of community populations. Second, our participants may not be representative of all Japanese community populations because our study was conducted only in a rural district. Even if the prevalence of depression does not differ across communities, certain risk factors could differ among different populations.³⁴

Conclusion

The CES-D exhibited good validity, and a score of 16 is the optimal cutoff score for assessment of MDD in a community population. Moreover, the use of SSLRs could be a convenient and intuitive measure for understanding the results

of community-based screening for depression. Programs combining sequential screening for depression and feedback with adequate support are recommended.

Acknowledgments

The authors thank all of their colleagues in this study for their skillful contributions to the collection and management of the data. Funding for this study was provided by the Hirosaki Research Institute for the Neurosciences. The Hirosaki Research Institute for the Neurosciences had no further role in study design; the collection, analysis, and interpretation of data; the writing of the report; or in the decision to submit the paper for publication.

Disclosure

The authors report no conflicts of interest in this work.

References

- Kessler RC, Bromet EJ. The epidemiology of depression across cultures. *Annu Rev Public Health*. 2013;34:119–138.
- Ishikawa H, Kawakami N, Kessler RC; World Mental Health Japan Survey Collaborators. Lifetime and 12-month prevalence, severity and unmet need for treatment of common mental disorders in Japan: results from the final dataset of World Mental Health Japan Survey. *Epidemiol Psychiatr Sci*. 2016;25(3):217–229.
- World Health Organization [webpage on the Internet]. *The Global Burden of Disease 2004 Update*. Available from: http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/. Accessed August 4, 2017.
- Barracough B, Bunch J, Nelson B, Sainsbury P. A hundred cases of suicide: clinical aspects. *Br J Psychiatry*. 1974;125:355–373.
- Isometsä E, Henriksson M, Marttunen M, et al. Mental disorders in young and middle aged men who commit suicide. *BMJ*. 1995; 310(6991):1366–1367.
- Brent DA, Perper JA, Moritz G, et al. Psychiatric risk factors for adolescent suicide: a case-control study. *J Am Acad Child Adolesc Psychiatry*. 1993;32(3):521–529.
- Grøholt B, Ekeberg O, Wichstrøm L, Haldorsen T. Young suicide attempters: a comparison between a clinical and an epidemiological sample. *J Am Acad Child Adolesc Psychiatry*. 2000;39(7): 868–875.
- Turvey CL, Conwell Y, Jones MP, et al. Risk factors for late-life suicide: a prospective, community-based study. *Am J Geriatr Psychiatry*. 2002; 10(4):398–406.
- Oyama H, Koida J, Sakashita T, Kudo K. Community-based prevention for suicide in elderly by depression screening and follow-up. *Community Ment Health J*. 2004;40(3):249–263.
- Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1:385–401.
- Shima S, Shikano T, Kitamura T. [New self-rating scales for depression.] *Clin Psychiatry*. 1985;27:717–723. Japanese.
- Iwata N, Saito K. Relationships of the Todai Health Index to the General Health Questionnaire and the Center for Epidemiologic Studies Depression Scale. *Nihon Eiseigaku Zasshi*. 1987;42(4):865–873.
- Iwata N, Saito K. Psychometric properties of the center for epidemiologic studies depression scale of Japanese workers. *Sangyo Igaku*. 1989;31(1):20–21.
- Schmitz N, Kruse J, Tress W. Application of stratum-specific likelihood ratios in mental health screening. *Soc Psychiatry Psychiatr Epidemiol*. 2000;35(8):375–379.
- Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ*. 2004;329(7458):168–169.
- Furukawa TA, Andrews G, Goldberg DP. Stratum-specific likelihood ratios of the general health questionnaire in the community: help-seeking and physical co-morbidity affect the test characteristics. *Psychol Med*. 2002;32(4):743–748.
- Wada K, Tanaka K, Theriault G, Moriyama M, Satoh T, Aizawa Y. Application of the stratum-specific likelihood ratio (SSLR) analysis to results of a depressive symptoms screening survey among Japanese workers. *Soc Psychiatry Psychiatr Epidemiol*. 2007;42(5):410–413.
- Yang HJ, Soong WT, Kuo PH, Chang HL, Chen WJ. Using the CES-D in a two-phase survey for depressive disorders among nonreferred adolescents in Taipei: a stratum-specific likelihood ratio analysis. *J Affect Disord*. 2004;82(3):419–430.
- Funahashi K, Takahashi I, Danjo K, Matsuzaka M, Umeda T, Nakaji S. Smoking habits and health-related quality of life in a rural Japanese population. *Qual Life Res*. 2011;20(2):199–204.
- Sugawara N, Yasui-Furukori N, Takahashi I, Matsuzaka M, Nakaji S. Age and gender differences in the factor structure of the Center for Epidemiological Studies Depression Scale among Japanese working individuals. *Compr Psychiatry*. 2015;56:272–278.
- Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59(suppl 20):22–33.
- Otsubo T, Tanaka K, Koda R, et al. Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry Clin Neurosci*. 2005;59(5):517–526.
- Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988; 240(4857):1285–1293.
- Peirce JC, Cornell RG. Integrating stratum-specific likelihood ratios with the analysis of ROC curves. *Med Decis Making*. 1993;13(2): 141–151.
- Furukawa T, Hirai T, Kitamura T, Takahashi K. Application of the Center for Epidemiologic Studies Depression Scale among first-visit psychiatric patients: a new approach to improve its performance. *J Affect Disord*. 1997;46(1):1–13.
- Simel DL, Samsa GP, Matchar DB. Likelihood ratios for continuous test results – making the clinicians' job easier or harder? *J Clin Epidemiol*. 1993;46(1):85–93.
- Stockings E, Degenhardt L, Lee YY, et al. Symptom screening scales for detecting major depressive disorder in children and adolescents: a systematic review and meta-analysis of reliability, validity and diagnostic utility. *J Affect Disord*. 2015;174:447–463.
- Vilagut G, Forero CG, Barbaglia G, Alonso J. Screening for depression in the general population with the center for epidemiologic studies depression (CES-D): a systematic review with meta-analysis. *PLoS One*. 2016;11(5):e0155431.
- Patten SB. Performance of the composite international diagnostic interview short form for major depression in community and clinical samples. *Chronic Dis Can*. 1997;18(3):109–112.
- Gupta S, Goren A, Dong P, Liu D. Prevalence, awareness, and burden of major depressive disorder in urban China. *Expert Rev Pharmacoecon Outcomes Res*. 2016;16(3):393–407.
- Thombs BD, Coyne JC, Cuijpers P, et al. Rethinking recommendations for screening for depression in primary care. *CMAJ*. 2012;184(4): 413–418.
- Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med*. 2008;5(2):e45.
- Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA*. 2010;303(1):47–53.
- Fujise N, Abe Y, Fukunaga R, et al. Comparisons of prevalence and related factors of depression in middle-aged adults between urban and rural populations in Japan. *J Affect Disord*. 2016;190:772–776.

Neuropsychiatric Disease and Treatment**Dovepress****Publish your work in this journal**

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS,

and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>