The Validity of the Montgomery-Asberg Depression Rating Scale in an Inpatient Sample with Alcohol Dependence

Breanne Hobden (B Psyc)[1,2,3], Melanie L. Schwandt (Ph.D)[4], Mariko Carey (Ph.D)[2,3], Mary R. Lee (M.D.)[1], Mehdi Farokhnia (M.D.)[1], Sofia Bouhlal (Ph.D)[1], Christopher Oldmeadow (Ph.D)[5,6] and Lorenzo Leggio (M.D., Ph.D)[1,7*]

1. Section on Clinical Psychoneuroendocrinology and Neuropsychopharmacology, National Institute on Alcohol Abuse and Alcoholism and National Institute on Drug Abuse; National Institutes of Health, Bethesda, MD, USA
2. Health Behaviour Research Group. HMRI Building, University of Newcastle, New South Wales 2308, Australia
3. Priority Research Centre in Health Behaviour and Hunter Medical Research Institute. HMRI Building, University of Newcastle, New South Wales 2308, Australia
4. Office of the Clinical Director, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, USA
5. Public Health Stream, Hunter Medical Research Institute. HMRI Building, New South Wales 2308, Australia
6. Centre for Clinical Epidemiology and Biostatistics, University of Newcastle. HMRI Building, University of Newcastle, New South Wales 2308, Australia
7. Center for Alcohol and Addiction Studies, Department of Behavioral and Social Sciences, Brown University, Providence, RI, USA

Word count: 3,751

Running title: Validity of the MADRS for inpatients with alcoholism

*Corresponding Author:

Lorenzo Leggio, M.D., Ph.D., M.Sc.

Section on Clinical Psychoneuroendocrinology and Neuropsychopharmacology

NIAAA & NIDA, NIH

10 Center Drive (10CRC/15330) MSC 1108; Room 1-5429, Bethesda, MD 20892-1108

Phone: +1 301 435 9398; Fax: + 1 301 402 0445; E-mail: lorenzo.leggio@nih.gov

Acknowledgements

This work was supported by National Institutes of Health (NIH) intramural funding ZIA-AA000218, *Section on Clinical Psychoneuroendocrinology and Neuropsychopharmacology* (CPN; PI: Leggio), jointly supported by the Division of Intramural Clinical and Biological Research of the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the Intramural Research Program of the National Institute on Drug Abuse (NIDA). Ms. Breanne Hobden's visiting fellowship in Dr. Leggio's CPN laboratory was supported by the Australian Academy of Science on behalf of the Adam J Berry Memorial Fund (AJBMF). Ms. Breanne Hobden is supported by an Australian Rotary Health, Ian Scott Mental Health PhD Scholarship. Dr. Mariko Carey is supported by a National Health and Medical Research Council (NHMRC) Translating Research into Practice Fellowship (G1300684).

The authors would like to thank the clinical and research staff involved in data collection and support at the NIAAA Division of Intramural Clinical and Biological Research, and at the NIH Clinical Center (*e.g.* Department of Nursing). The authors would also like to thank Ms. Karen Smith, NIH Library for bibliographic assistance.

Disclosures

The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors report no biomedical financial interests or potential conflicts of interest.

Abstract (300 words)

Background: The Montgomery-Asberg Depression Rating Scale (MADRS) is commonly used to examine depressive symptoms in clinical settings, including facilities treating patients for alcohol addiction. No studies have examined the validity of the MADRS compared to an established clinical diagnostic tool of depression in this population. This study aimed to examine: 1) the validity of the MADRS compared to a clinical diagnosis of a depressive disorder (using the Structured Clinical Interview for DSM-IV (SCID)) in patients seeking treatment for alcohol dependence (AD); 2) whether the validity of the MADRS differs by type of SCID-based diagnosis of depression; and 3) which items contribute to the optimal predictive model of the MADRS compared to a SCID diagnosis of a depressive disorder.

Methods: Individuals seeking treatment for AD and admitted to an inpatient unit were administered the MADRS at day 2 of their detoxification program. Clinical diagnoses of AD and depression were made via the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders-IV at the beginning of treatment.

Results: In total, 803 participants were included in the study. The MADRS demonstrated low overall accuracy relative to the clinical diagnosis of depression with an area under the curve of 0.68. The optimal threshold for balancing sensitivity and specificity identified by the Euclidean distance was >14. This cut-point demonstrated a sensitivity of 66%, a specificity of 60%, a positive predictive value of 50% and a negative predictive value of 75%. The MADRS performed slightly better for major depressive disorders compared to alcohol-induced depression. Items related to lassitude, concentration and appetite slightly decreased the accuracy of the MADRS.

Conclusion: The MADRS does not appear to be an appropriate substitute for a diagnostic tool among alcohol-dependent patients. The MADRS may, however, still be a useful screening tool assuming careful consideration of cut-off scores.

Key words: Alcohol-Related Disorders, Depression, Dual Diagnosis, Sensitivity and Specificity, ROC Curve

Introduction

Among individuals with mood disorders, approximately 22% have a comorbid substance use disorder (Conway et al., 2006), while 25% of individuals with addictive disorders report mood disorders within the past year (Kessler et al., 1996). These data suggest a high degree of co-morbidity between these two mental health problems. Indeed, a recent meta-analysis of epidemiological surveys examining comorbid substance use, mood disorders and anxiety found that mood disorders were three times more prevalent among those with alcohol dependence (AD) (Lai et al., 2015). Additionally, increased depressive and AD symptoms severity (Burns et al., 2005, Sullivan et al., 2005) increases the likelihood of seeking treatment for addiction, resulting in higher rates of depression within populations admitted to addiction treatment facilities (Tolliver and Anton, 2015, Kodl et al., 2007). High quality care for these individuals relies on accurate diagnosis of depressive symptoms to establish the optimal course of treatment for both depression and addiction.

The Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders (SCID) is considered one of the gold standard examinations to diagnose depression (Cohen, 1998). However, a SCID administration is time consuming for patients and clinicians and requires extensive administrator training (Biometrics Research Department, Cohen, 1998). Therefore, in clinical settings and research studies, there is an increased tendency to employ brief screening questionnaires for detecting depressive symptoms (Henkel et al., 2004, Mitchell et al., 2012). While brief measures are useful tools to minimize the burden of administration, their usefulness is determined by their accuracy compared to gold standard assessment via clinical interview. The Montgomery-Asberg Depression Rating Scale (MADRS) is one of the most commonly used depression measures in research and clinical settings (Behzadifar et al., 2015, Mrazek et al., 2014). The MADRS has been found to have strong psychometric properties among patients with depressive disorders (Hawley et al.,

2002, Williams and Kobak, 2008) and has shown to discriminate between depression severity levels (Muller et al., 2000, Muller et al., 2003). The MADRS was specifically developed to be sensitive to changes in depressive symptoms over time, making it a particularly useful tool for monitoring patients undergoing treatment and participants involved in clinical trials (Montgomery and Asberg, 1979).

The MADRS has minimal focus on querying somatic symptoms. Therefore, it may be useful in identifying depression in the AD population where somatic comorbidity is common. This has previously been demonstrated in other populations where somatic overlap of symptoms occurs, *e.g.* bariatric surgery patients (Duarte-Guerra et al., 2016), patients with Parkinson's disease (Leentjens et al., 2000) and geriatric populations with ovarian cancer (Rhondali et al., 2015). These studies found the MADRS to have a high sensitivity and specificity compared to a clinical diagnosis of a depressive disorder. However, given the potential for symptom overlap between depression and AD, some MADRS items, such as those focussing on appetite loss or concentration difficulties, may decrease the accuracy of this scale in determining depression. It is therefore useful to examine which items contribute most strongly to the accurate detection of depression in patients with AD. To date, the short versions of the MADRS have not been validated in an AD population.

Due to the overlap of symptoms and the bidirectional relationship between AD and depression, it is often difficult for health professionals to differentiate between depression induced by AD (henceforth referred to as alcohol-induced depression) and primary depression (*e.g.* major depression) (Tolliver and Anton, 2015). Current treatment approaches for differentiating between these two categories of depression among AD inpatients involves patients undergoing a period of abstinence to determine whether the depressive symptoms remain or subside (Dongier, 2005). As pharmacological treatment for patients with co-morbid depression and AD has been associated with improved AD outcomes (Pettinati et al., 2010),

providing treatment during the early stages of AD treatment may increase the likelihood of successful outcomes. However, it has also been suggested that as alcohol-induced depression is a consequence of AD symptoms, depressive symptoms may subside with abstinence rendering the use of medication unnecessary, costly and burdensome (Pettinati, 2004). Therefore, being able to differentiate between alcohol-induced and primary depression may assist clinicians in determining the optimal treatment approach.. Therefore, there is value in exploring the effectiveness of the MADRS in detecting both depression types in AD population.

While the MADRS is commonly used to examine depression among those with AD (Muhonen et al., 2011, Muhonen et al., 2008, Gual et al., 2003), no studies have assessed its validity as a diagnostic tool for depression compared to a gold standard diagnostic tool, such as the SCID, in these patients. This study therefore aimed to examine 1) the validity of the MADRS among an inpatient group seeking treatment for AD through exploring its sensitivity, specificity, positive and negative predictive power at different thresholds compared to a SCID diagnosis of a depressive disorder; 2) whether the validity of the MADRS differs by type of SCID-based diagnosis of depression (*i.e.* alcohol-induced versus major depression); and 3) which items contribute to the optimal predictive model of the MADRS compared to a SCID diagnosis of a depressive disorder.

Methods

Participants and Procedures

The data for this study were extracted from a larger database held by the National Institute on Alcohol Abuse and Alcoholism (NIAAA). This database included a sample of individuals seeking treatment for AD and admitted to an inpatient unit at the National

Institutes of Health (NIH) Clinical Center for an NIAAA alcohol detoxification program. The inpatient detoxification period lasted approximately 30 days.

Participants were recruited from December 2006 to June 2016 through physicians' referrals, word of mouth, community outreach, NIH websites, and online and newspaper advertisements. Participants were evaluated and received patient care under the NIAAA screening protocols approved by the appropriate NIH Institutional Review Boards. Individuals who were phone-screened for potential participation to the alcohol detoxification program were provided with relevant information on the program. Those interested and eligible were scheduled for inpatient admission. After they signed a written consent form, they were administered a battery of screening tests. Further assessments were administered on day 2 and throughout the remainder of the inpatient stay. The MADRS was administered on the 2nd day of the inpatient detoxification period via interview by trained clinical staff. The SCID interview was administered approximately 10 days after admission.

Inclusion criteria for this analysis were a current diagnosis of AD according to the DSM for Mental Disorders, 4th Edition, Text-revised (DSM-IV-TR) and available baseline MADRS data (measured on day 2 of admission). Lifetime diagnosis of bipolar disorder, schizophrenia or other psychotic disorders were exclusion criteria for this analysis.

Main Assessments/Measures

Clinical psychiatric diagnosis: The SCID (First et al., 2002) was used for diagnosing all axis 1 disorders including AD and depressive disorders. Depressive disorders included the following: alcohol induced-mood disorders, major depressive disorders (recurrent, single episode and unspecified), dysthymic disorders, medical mood disorders, current bereavement and depressive disorders not otherwise specified. The timeframe for current disorders included a cluster of symptoms present during the same 2-week period occurring within the

past month. Henceforth, a DSM diagnosis of depression refers to a diagnosis based on the SCID.

Depressive Symptoms: The MADRS, a 10-item scale (range: 0-60), was used to examine scores for depressive symptoms over the past week (*e.g.* reported sadness, inner tension, etc.)(Montgomery and Asberg, 1979). Previous studies have recommended the following severity estimates based on the MADRS score: 0 to 6 = no depression; 7 to 19 = mild depression; 20 to 34 = moderate depression; >34 = severe depression (Snaith et al., 1986, Herrmann et al., 1998).

Demographic characteristics: Gender, age, years of education and race were collected for all participants during screening.

Additional Assessments/Measures

The following clinical and research assessments/measures were collected during the inpatient detoxification period and were used for this analysis:

Alcohol drinking: A 90-day Timeline Follow-Back (TLFB) (Sobell and Sobell, 1992) questionnaire was used to determine alcohol consumption prior to admission. The TLFB is a semi-structured interview aimed at estimating daily alcohol consumption. Several outcome measures can be inferred from the TLFB, including: total drinks, number of drinking days, number of heavy drinking days, and average number of drinks per drinking day.

Alcohol Dependence Severity (ADS): The ADS is a 25 item self-report scale (range: 0-47) used to measure the severity of AD (Skinner and Allen, 1982).

Clinical Institute Withdrawal Assessment of Alcohol Scale, Revised (CIWA-Ar): To evaluate the severity of alcohol withdrawal and if necessary, its appropriate medical treatment, the 10-item CIWA-Ar (range: 0-67) (Sullivan et al., 1989) was administered approximately every 2

to 4 hours or according to clinical judgement, for approximately the first week of admission. An overall maximum CIWA-Ar score was calculated using the highest CIWA-Ar measurement taken across the seven days. Benzodiazepine dose was recorded by clinicians each time it was administered.

Statistical analysis

Descriptive statistics were used to evaluate patients' characteristics. Comparisons of the characteristics between those with and without a diagnosis of any DSM depressive disorder were performed using an independent two-sample t-test for continuous variables and Chi-squared test for categorical variables. To assess the performance of the MADRS at baseline at predicting a SCID diagnosis of a depressive disorder, empirical receiver operating characteristics (ROC) curves were constructed using estimates of sensitivity and 1-specificity for each cut-point. Positive predictive values (PPV) and negative predictive values (NPV) were also estimated. The area under the ROC curve (AUC) was then estimated and categorized as either having low accuracy (>0.5 and <0.7), moderate accuracy (≥0.7 and <0.9), or high accuracy (≥0.9) (Cairney et al., 2007). The minimum Euclidean distance was used to define the point on the ROC curve that is closest to a perfect predictor (i.e. sensitivity of 100% and a false positive rate of zero). The sample with a SCID diagnosis of a depressive disorder were then split into alcohol-induced and primary major depressive disorders. ROC curves were applied to each of these groups to determine if the type of diagnosis impacted the accuracy of the MADRS compared to the SCID diagnosis. There were insufficient numbers to assess other categories of depressive disorders (*e.g.* dysthymia). Lastly, we constructed ROC curves for each individual item of the MADRS among the entire sample, the alcohol-induced depression group and the major depressive disorder group. A series of univariate

logistic regression models were used to assess the predictive performance of each item on the MADRS. Items were ranked by their Akaike Information Criterion (AIC) to determine which items contribute the most to the optimal overall model of the MADRS compared to a SCID diagnosis of a depressive disorder. AIC is a better measure for model comparison than AUC; it can be thought of as an estimate of the out of sample predictive error. Separate multi-variate models were fit with increasing number of items. The AUC for the model with the lowest AIC is reported. Higher ranking items were combined to determine which combination of items provided an optimal AUC and to allow the removal of any redundant items. The alpha level for determining statistical significance was set at 0.05. All statistical analyses were conducted with SAS version 9.2 (SAS Institute, Cary, NC).

Results

Sample

A total of 803 participants met the inclusion criteria and were included in the analysis. Sample characteristics are reported in **Table 1**. Briefly, the overall sample included 571 males (71.1%), had an average age of 43.0 years (Standard Deviation (SD)=10.5) and the predominant race was Caucasian (*n* = 423; 52.7%). In addition to AD, 42.4% of the sample had one or more DSM diagnoses of current dependence for another substance.

<<Insert Table 1 here>>

MADRS accuracy for detecting depression among inpatients with AD

There were 302 (37.6%) participants with a current DSM-IV-TR diagnosis of a depressive disorder identified by the SCID. A comparison of characteristics between those with and those without a diagnosis of a depressive disorder can be found in **Table 1**. Sensitivity, specificity, PPV and NPV for MADRS scores ranging from 7 to 26 can be found in **Table 2**. The MADRS demonstrated low overall accuracy relative to a clinical diagnosis for discriminating between those with and those without a SCID diagnosis of a depressive disorder. The AUC was 0.68, which was statistically significant, $\chi 2$ (df = 1, N = 803) = 65.07, $P < 0.0001$ (**Figure 1**). The optimal threshold for balancing sensitivity and specificity identified by the minimum Euclidean distance was >14. At this cut-point, the MADRS correctly identified 66% of depression cases (sensitivity) and 60% of non-cases for depression (specificity). Only 50% of cases identified as depression by the MADRS, using the >14 cut-point, were classified as such by the SCID diagnosis (PPV), while 75% of patients who were identified by the MADRS (score of ≤14) as non-cases of depression were classified as such according to the SCID (NPV).

<<Insert Table 2 here>>

<<Insert Figure 1 here>>

SCID diagnosis of alcohol-induced depression

Among the clinically depressed patients, 167 (55.3%) had a SCID diagnosis of alcohol-induced depression. We applied a ROC curve to this subgroup (*Figure S1 and Table S1*). The MADRS demonstrated low overall accuracy with an AUC of 0.64. The optimal threshold for

balancing sensitivity and specificity identified by an Euclidean distance optimal threshold of >14, the same threshold identified for the overall sample. This cut-point demonstrated a sensitivity of 62%, a PPV of 33%, a specificity of 60% and a NPV of 83%.

SCID diagnosis of a major depressive disorder

Among the clinically depressed patients, 82 (27.2%) of the sample had a SCID diagnosis of a major depressive disorder. The ROC curve demonstrated an AUC of 0.73 (*Figure S2 and Table S2*). The optimal threshold for balancing sensitivity and specificity identified by an Euclidean distance optimal threshold of >18, slightly higher than that observed for the overall sample. This cut-point demonstrated a sensitivity of 61%, a PPV of 29%, a specificity of 76% and a NPV of 92%.

ROC curve analyses of individual MADRS items

The ROC curve for each individual item of the MADRS among the entire sample, the alcohol-induced depression group and the major depressive disorder group revealed that in all three models the items: “lassitude”, “concentration difficulties” and “reduced appetite” slightly decreased the MADRS accuracy. When the overall model was run without these three items, the AUC increased to 0.69 with an Euclidean distance optimal threshold of 11. This cut-point demonstrated a sensitivity of 61%, a PPV of 52%, a specificity of 66% and a NPV of 74%. For the alcohol-induced depression group, “pessimistic thoughts” also slightly decreased the accuracy of the MADRS. When “lassitude”, “concentration difficulties”, “reduced appetite” and “pessimistic thoughts” were removed from the model, the AUC was 0.64 with an Euclidean distance optimal threshold of 8. This cut-point demonstrated a sensitivity of 65%, a PPV of 34%, a specificity of 58% and a NPV of 83%. In the major depressive disorder group inner tension slightly decreased the AUC. When “lassitude”, “concentration difficulties”, “reduced appetite” and “inner tension” were removed from the

model, the AUC was 0.75 with an Euclidean distance optimal threshold of 10. This cut-point demonstrated a sensitivity of 70%, a PPV of 28%, a specificity of 71% and a NPV of 93%.

Discussion

To our knowledge, this study is the first to examine the validity of the MADRS in an alcohol-dependent sample. The results of this study indicate that the MADRS does not have strong predictive capabilities for balancing sensitivity and specificity of a depressive diagnosis among alcohol-dependent individuals recently hospitalized in an inpatient detoxification setting. The ROC curve analysis demonstrated a low AUC with the optimal cut-point demonstrating a high rate of false positives and negatives. The MADRS has shown good discriminate properties between those with and without a DSM diagnosis of depression among other clinical settings. Duarte-Guerra and colleagues found a 13/14 cut score on the MADRS demonstrated a sensitivity of 85% and specificity of 81% among bariatric surgery patients (Duarte-Guerra et al., 2016). Similarly, Leentjens *et al.* found a 14/15 cut score among patients with Parkinson's Disease had a sensitivity of 88% and a specificity of 89% (Leentjens et al., 2000), and Rhondali *et al.* found a cut score of 16 to have a 88% sensitivity and 91% specificity in elderly patients with ovarian cancer (Rhondali et al., 2015). The present study failed to replicate these findings in an alcohol-dependent inpatient sample. The optimal cut-point identified by the Euclidean distance was 14; however, as this cut-point had a relatively low sensitivity and specificity, we are unable to endorse the use of a cut-point for identifying potential depression using the MADRS. This finding may be due to AD and depression both being mental health conditions rather than one being a physical condition, as was the case with the previous studies (Leentjens et al., 2000, Rhondali et al., 2015, Duarte-Guerra et al., 2016). Distinguishing between the two conditions may be more difficult among our study sample due the possibility of alcohol-induced depressive symptoms. It is

conceivable that these symptoms would have been examined more thoroughly during the SCID interview compared to the MADRS and this could have resulted in different interpretations of these symptoms, contributing to discrepancies between a clinical diagnosis of depression and the MADRS scores. Additionally, this finding may indicate that the MADRS alone is not enough to measure depression in an AD sample due to the multilayered and complex nature of addiction. Irrespectively, both the MADRS and a DSM diagnosis theoretically measure the same construct (*i.e.* depression), therefore one would expect greater convergence between these measures (Bagozzi et al., 1991). The findings from this study lead to question the construct validity of the MADRS among this specific population, *i.e.* alcohol-dependent individuals.

While the MADRS did not demonstrate strong properties as a diagnostic tool among this sample, the cut-points may still prove to be useful for screening in different settings. For example, settings which have adequate resources to conduct follow-up diagnostic interviews can allow for a higher number of false positive results. In these settings a lower cut-point (*i.e.* >6 or >7) could be used to reduce the number of diagnostic interviews required while maximizing sensitivity. In settings where resources are scarce or where false positives need to be minimized, for instance when recruiting participants for a research study, a higher cut-point (*i.e.* >19 or >20) could be used. When examining the ROC curves among alcohol-induced depression and major depressive disorders, a slightly higher overall AUC was found for the group with a major depressive disorder. This may indicate that the MADRS is a better measure of depression when it is independent from AD as opposed to depression that may be secondary to AD. While the AUC was slightly higher for this group, in terms of the cut-point for optimizing sensitivity and specificity, this improvement was mostly exhibited through an increase in specificity, where sensitivity remained low.

Examination of the individual items of the MADRS demonstrated the items "lassitude", "concentration difficulties" and "reduced appetite" were associated with a decrease in AUC for the overall model. This finding may be due to these somatic symptoms potentially being related to patients' AD. While the overall change in AUC was not large after removing these items, the fact that the AUC did increase shows that these items could potentially be removed from the MADRS without impacting its validity in this population, thus decreasing burden on clinicians or researchers administering the tool. However, this speculation needs to be further tested in order to directly assess the potential validity of such modified MADRS.

This study should be seen in light of its strengths and limitations. This study has one of the largest sample sizes used to evaluate MADRS in a targeted sample (Leentjens et al., 2000, Duarte-Guerra et al., 2016, Rhondali et al., 2015). The inpatient setting allowed for a careful monitoring of alcohol abstinence and withdrawal. Limitations include the difficulties associated with diagnosing depression in an alcohol-dependent population during the early phase of detoxification, particularly when differentiating alcohol-induced and non-alcohol induced depression. It is important to note, however, that when exploring these groups separately and together there were no significant changes in the AUC, sensitivity or specificity for the MADRS scores. Further limitations were the difference in length of time over which symptoms were assessed (1 week for MADRS, 1 month for SCID) and the difference in administration time between the SCID interview (approximately 10 days after admission) and the MADRS (day 2 of admission). While these factors may have caused some discrepancy between the two measures, this is likely to have been moderately offset through both tools accounting for symptoms within a recent timeframe. Of note, such limitations are common in studies of this kind, as previously reported (Gjerdingen et al., 2011). In general, it is possible that analyzing the MADRS later during the inpatient stay and after the resolution

of withdrawal symptoms may yield different results, *i.e.* an improved accuracy of the measure. However, we tested this hypothesis in our cohort by looking at MADRS assessments performed later during the inpatient stay: the accuracy of the MADRS was not improved but the overall cut points were lower because scores had reduced during the alcohol detoxification (*data not shown*).

Future research could replicate this work in a different setting, such as among alcohol-dependent patients seeking treatment for AD in an outpatient setting and/or among individuals seeking treatment for depression with comorbid AD to examine this group as an intermediate phenotype. Furthermore, while our sample included patients diagnosed from 2006 to 2016 via the DSM-IV, future work is needed to replicate this work in patients with the recently implemented DSM-5. Finally, future research should focus on comparing the MADRS and other tools to one another, specifically in an alcohol-dependent population. For example, previous research has demonstrated that the Patient Health Questionnaire, a 9-item self-administered measure based upon the diagnostic criteria of the DSM, has good sensitivity and specificity (Delgadillo et al., 2011) and strong psychometric properties (Dum et al., 2008) in a substance abuse setting. The Beck Depression Inventory, a 21-item self-administered measure typically used to gauge depression severity, also has good psychometric properties among alcohol and other drugs users (McPherson and Martin, 2010, Dum et al., 2008).

In conclusion, the results of this study indicate the MADRS administered early at admission may not be a suitable tool for determining the presence of a depressive disorder in AD inpatient populations, when conducting a full SCID interview is not possible. The lack of convergence between the MADRS scores and a SCID-based DSM-IV diagnosis of depression highlights a potential lack of construct validity of the MADRS in this population. While the MADRS may still be useful as a screening tool to minimize the number of

diagnostic interviews required, the findings from this study have significant implications for use of the MADRS in gauging depressive symptoms at the beginning of alcohol treatment and for determining eligibility in clinical trials. Clinicians and researchers should carefully consider the strengths and weaknesses of this tool before employing it in alcohol-dependent patients.

References

Bagozzi, R. P., Yi, Y. & Phillips, L. W. 1991. Assessing construct validity in organizational research. *Administrative Science Quarterly,* 36**,** 421-458.

Behzadifar, M., Dehghan, H., Saki, K., Behzadifar, M., Keshavarzi, A., Saran, M. & Sari, A. A. 2015. Evaluation efficacy and safety of Vortioxetine 20 mg/d versus placebo for treatment Major Depressive Disorder: A systematic review and meta-analysis of randomized controlled trials. *Pharmacology & Pharmacy,* 6**,** 221-231.

Biometrics Research Department. *SCID training sequence of steps* [Online]. Available: http://www.scid4.org/training/overview.html [Accessed 5th Oct 2016].

Burns, L., Teesson, M. & O'neill, K. 2005. The impact of comorbid anxiety and depression on alcohol treatment outcomes. *Addiction,* 100**,** 787-796.

Cairney, J., Veldhuizen, S., Wade, T. J., Kurdyak, P. & Streinter, D. L. 2007. Evaluation of 2 measures of psychological distress as screeners for depression in the general population. *Can J Psychiatry,* 52**,** 111-120.

Cohen, S. 1998. *Measures of Depression as a Clinical Disorder* [Online]. The Psychosocial Working Group. Available: http://www.macses.ucsf.edu/research/psychosocial/depression.php [Accessed 12 October 2016].

Conway, K. P., Compton, W., Stinson, F. S. & Grant, B. F. 2006. Lifetime comorbidity of DSM-IV mood and anxiety disorders and specific drug use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry,* 67**,** 247-57.

Delgadillo, J., Payne, S., Gilbody, S., Godfrey, C., Gore, S., Jessop, D. & Dale, V. 2011. How reliable is depression screening in alcohol and drug users? A validation of brief and ultra-brief questionnaires. *Journal of Affective Disorders,* 134**,** 266-271.

Dongier, M. 2005. What are the treatment options for comorbid alcohol abuse and depressive disorders? *J Psychiatry Neurosci,* 30**,** 224.

Duarte-Guerra, L. S., Gorenstein, C., Paiva-Medeiros, P. F., Santo, M. A., Neto, F. L. & Wang, Y.-P. 2016. Clinical utility of the Montgomery-Asberg Depression Rating Scale for the detection of depression among bariatric surgery candidates. *BMC Psychiatry,* 16**,** 119.

Dum, M., Pickren, J., Sobell, L. & Sobell, M. B. 2008. Comparing the BDI-II and the PHQ-9 with outpatient substance abusers. *Addictive behaviours,* 33**,** 381-387.

First, M. B., Spitzer, R. L., Miriam, G. & Williams, J. B. W. 2002. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P) New York: Biometrics Research, New York State Psychiatric Institute.

Gjerdingen, D., Mcgovern, P. & Center, B. 2011. Problems with a diagnostic depression interview in a postpartum depression trial. *Journal of the American Board of Family Medicine,* 24**,** 187-193.

Gual, A., Balcells, M., Torres, M., Madrigal, M., Diez, T. & Serrano, L. 2003. Sertraline for the prevention of relapse in detoxicated alcohol dependent patients with a comorbid depressive disorder: A randomized controlled trial. *Alcohol and Alcoholism,* 38**,** 619-625.

Hawley, C. J., Gale, T. M. & Sivakumaran, T. 2002. Defining remission by cut off score on the MADRS: selecting the optimal value. *Journal of Affective Disorders,* 72**,** 177-184.

Henkel, V., Mergl, R., Kohnen, R., Allgaier, A.-K., Moller, H.-J. & Hegerl, U. 2004. Use of brief depression screening tools in primary care: consideration of heterogeneity in performance in different patient groups. *General Hospital Psychiatry,* 26**,** 190-198.

Herrmann, N., Black, S., Lawrence, J., Szekely, C. & Szalai, J. 1998. The Sunnybrook stroke study: A prospective study of depressive symptoms and functional outcomes. *Stroke,* 29**,** 618-624.

Kessler, R. C., Nelson, C. B., Mcgonagle, K. A., Edlund, M. J., Frank, R. G. & Leaf, P. J. 1996. The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. *Am J Orthopsychiatry,* 66**,** 17-31.

Kodl, M. M., Fu, S. S., Willenbring, M. L., Gravely, A., Nelson, D. B. & Joseph, A. M. 2007. The Impact of Depressive Symptoms on Alcohol and Cigarette Consumption Following Treatment for Alcohol and Nicotine Dependence. *Alcoholism: Clinical and Experimental Research,* 32**,** 92-99.

Lai, H. M., Cleary, M., Sitharthan, T. & Hunt, G. E. 2015. Prevalence of comorbid substance use, anxiety and mood disorders in epidemiological surveys, 1990-2014: A systematic review and meta-analysis. *Drug Alcohol Depend,* 154**,** 1-13.

Leentjens, A. F. G., Verhey, F. R. J., Lousberg, R., Spitsbergen, H. & Wilmink, F. W. 2000. The validity of the Hamilton and Montgomery-Asberg Depression Rating Scales as screening and diagnostic tools for depression in Parkinson's Disease. *International Journal of Geriatric Psychiatry,* 15**,** 644-649.

Mcpherson, A. & Martin, C. R. 2010. A narrative review of the Beck Depression Inventory (BDI) and implications for its use in an alcohol-dependent population. *Journal of Psychiatric and Mental Health Nursing,* 17**,** 19-30.

Mitchell, A. J., Meader, N., Davies, E., Clover, K., Carter, G. L., Loscalzo, M. J., Linden, W., Grassi, L., Johansen, C., Carlson, L. E. & Zaboraemail, J. 2012. Meta-analysis of screening and case finding tools for depression in cancer: Evidence based recommendations for clinical practice on behalf of the Depression in Cancer Care consensus group. *Journal of Affective Disorders,* 140**,** 149-160.

475 Montgomery, S. A. & Asberg, M. 1979. A new depression scale designed to be sensitive to
476 change. *British Journal of Psychiatry,* 134**,** 182-18.
477 Mrazek, D. A., Hornberger, J. C., Altar, C. A. & Degtiar, I. 2014. A review of the clinical,
478 economic, and societal burden of treatment-resistant depression: 1996-2013.
479 *Psychiatric Services,* 65**,** 977-987.
480 Muhonen, L. H., Lahti, J., Alho, H., Lönnqvist, J., Haukka, J. & Saarikoski, S. T. 2011.
481 Serotonin transporter polymorphism as a predictor for escitalopram treatment of
482 major depressive disorder comorbid with alcohol dependence. *Psychiatry Research,*
483 186**,** 53-57.
484 Muhonen, L. H., Lonngvist, J., Juva, K. & Alho, H. 2008. Double-blind, randomized
485 comparison of memantine and escitalopram for the treatment of major depressive
486 disorder comorbid with alcohol dependence. *Journal of Clincial Psychiatry,* 69**,** 392-
487 399.
488 Muller, M. J., Himmerich, H., Kienzl, B. & Szegedi, A. 2003. Differentiating moderate and
489 severe depression using the Montgomery–Asberg depression rating scale (MADRS).
490 *Journal of Affective Disorders,* 77**,** 255-260.
491 Muller, M. J., Szegedi, A., Wetzel, H. & Benkert, O. 2000. Moderate and severe depression
492 gradations for the Montgomery-Asberg Depression Rating Scale. *Journal of Affective*
493 *Disorders,* 69**,** 137-140.
494 Pettinati, H. 2004. Antidepressant treatment of co-occurring depression and alcohol
495 dependence. *Biological Psychiatry,* 56**,** 785-792.
496 Pettinati, H. M., Oslin, D. W., Kampman, K. M., Dundon, W. D., Xie, H., Gallis, T. L.,
497 Dackis, C. A. & O'brien, C. P. 2010. A double-blind, placebo-controlled trial
498 combining Sertraline and Natlrexone for treating co-occurring depression and alcohol
499 dependence. *American Journal of Psychiatry,* 167**,** 668-675.

Rhondali, W., Freyer, G., Adam, V., Filbet, M., Derzelle, M., Abgrall-Barbry, G., Bourcelot, S., Machavoine, J., Chomat-Neyraud, M., Gisserot, O., Largillier, R., Le Rol, A., Priou, F., Saltel, P. & Falandry, C. 2015. Agreement for depression diagnosis between DSM-IV-TR criteria, three validated scales, oncologist assessment, and psychiatric clinical interview in elderly patients with advanced ovarian cancer. *Clinical Interventions in Aging,* 10**,** 1155-1162.

Skinner, H. A. & Allen, B. A. 1982. Alcohol dependence syndrome: measurement and validation. *Journal of Abnormal Psychology,* 91**,** 199-209.

Snaith, R., Harrop, F., Newby, D. & Teale, C. 1986. Grade scores of the Montgomery-Asberg Depression and the Clinical Anxiety Scales. *British Journal of Psychiatry,* 148**,** 599-601.

Sobell, L. C. & Sobell, M. B. 1992. Timeline Follow-back: A technique for assessing self-reported ethanol consumption. *In:* ALLEN, J. & LITTEN, R. Z. (eds.) *Measuring Alcohol Consumption: Psychosocial and Biological Methods* Totowa, NJ: Humana Press.

Sullivan, J. T., Sykora, K., Schneiderman, J., Naranjo, C. A. & Sellers, E. M. 1989. Assessment of alcohol withdrawal: the revised clinical institute withdrawal assessment for alcohol scale (CIWA-Ar). *British Journal of Addiction,* 84**,** 1353-1357.

Sullivan, L. E., Fiellin, D. A. & O'connor, P. G. 2005. The prevalence and impact of alcohol problems in major depression: A systematic review. *American Journal of Medicine,* 118**,** 330-341.

Tolliver, B. K. & Anton, R. F. 2015. Assessment and treatment of mood disorders in the context of substance abuse. *Dialogues Clin Neurosci,* 17**,** 181-190.

523 Williams, J. B. W. & Kobak, K. A. 2008. Development and reliability of a structured
524 interview guide for the Montgomery-Asberg Depression Rating Scale (SIGMA).
525 *British Journal of Psychiatry,* 192**,** 52-58.

526

Legend to Figure

Figure 1. Receiver Operating Characteristic curve demonstrating an area under of the curve of 0.68 for the 10-item Montgomery-Asberg Depression Rating Scale scores compared to Structured Clinical Interview for a DSM diagnosis of a depressive disorder, with an optimum threshold of 14 using the Euclidean distance.

533 **Table 1.** Comparison of demographics and characteristics between those with and without a

534 Structured Clinical Interview for a DSM diagnosis of a depressive disorder.

Measure	**Positive Depressive Disorder Diagnosis**	**No Depressive Disorder Diagnosis**	**Overall Sample**	**Between Group *P*-Value**
Number of subjects	302 (37.6%)	501 (62.4%)	803 (100%)	.
Gender: *n* (%)				
Males	195 (64.6%)	376 (75.0%)	232 (28.9%)	$P = 0.0017$*
Females	107 (35.4%)	125 (25.0%)	571 (71.1%)	.
Age (years): M (SD)	43.1 (±10.3)	43.0 (±10.6)	43.0 (±10.5)	$P = 0.8619$
Education (years): M (SD)	13.6 (±2.5)	13.6 (±2.7)	13.6 (±2.6)	$P = 0.8861$
Race^: *n* (%)				
Caucasian	170 (56.3%)	253 (50.5%)	423 (52.7%)	$P = 0.2202$
African-American	105 (34.8%)	205 (40.9%)	310 (38.6%)	.
Asian	4 (1.3%)	5 (1.6%)	9 (1.1%)	.
American Indian/Alaskan	1 (0.3%)	3 (0.6%)	4 (0.5%)	.
Multiracial	9 (3.0%)	6 (1.2%)	1 (0.1%)	.
Native Hawaiian/Pacific Islander	0 (0%)	1 (0.2%)	41 (5.1%)	.
Unknown	13 (4.3%)	28 (5.6%)	15 (1.9%)	.
Alcohol Dependence Severity: M (SD)	23.2 (±7.8)	19.8 (±8.0)	21.1 (±8.1)	$P < 0.0001$*
Timeline Follow-back (last 90 Days): M (SD)				
Total Drinks	1068.1 (±760.9)	1041.1 (±727.2)	1051.3 (±739.7)	$P = 0.6237$

Number of Drinking Days	70.6 (±22.5)	71.1 (±22.5)	70.9 (±22.5)	*P* = 0.7545
Number of Heavy Drinking Days	66.1 (±25.2)	65.1 (±27.2)	65.5 (±26.5)	*P* = 0.6226
Average number of Drinks per Drinking Day	14.9 (±8.6)	14.0 (±8.2)	14.3 (±8.3)	*P* = 0.1820
Average total dose of benzodiazepines (mg) administered during inpatient stay: M (SD)	53.6 (±87.1)	35.2 (±69.7)	42.2 (±77.2)	*P* = 0.0019*
Overall Max CIWA-Ar: M (SD)	9.4 (±6.5)	7.3 (±6.0)	8.1 (±6.3)	*P* < 0.0001*
Other Current Substance Dependence: *n* (%)	141 (46.7%)	199 (39.7%)	346 (42.4%)	*P* = 0.053

535 *CIWA-Ar= Clinical Institute Withdrawal Assessment of Alcohol Scale, Revised; N=Number;*

536 *M=Mean; SD=Standard Deviation.*

537 ^Fisher's Exact test used due to low cell count.

538 *Statistically significant *P*<0.05

539

Table 2. Sensitivity, specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) for a range of Montgomery-Asberg Depression Rating Scale (MADRS) cut-off scores compared to Structured Clinical Interview for a DSM diagnosis of a depressive disorder.

MADRS Score	All Diagnoses of Depression*			
	Sensitivity	Specificity	PPV	NPV
>6	0.91	0.24	0.42	0.81
>7	0.88	0.27	0.42	0.80
>8	0.85	0.33	0.43	0.79
>9	0.82	0.39	0.45	0.79
>10	0.79	0.43	0.46	0.78
>11	0.76	0.47	0.46	0.77
>12	0.74	0.52	0.48	0.77
>13	0.70	0.56	0.49	0.75
>14	0.66	0.60	0.50	0.74
>15	0.61	0.64	0.51	0.73
>16	0.58	0.67	0.52	0.73
>17	0.55	0.72	0.54	0.72
>18	0.48	0.76	0.54	0.71
>19	0.44	0.78	0.55	0.70
>20	0.41	0.81	0.56	0.70
>21	0.38	0.83	0.57	0.69
>22	0.34	0.85	0.58	0.68
>23	0.31	0.87	0.58	0.68

>24	0.29	0.88	0.60	0.67
>25	0.26	0.90	0.62	0.67
>26	0.23	0.92	0.64	0.67

544 **Depression diagnoses included: alcohol induced- mood disorders, major depressive*

545 *disorders (recurrent, single episode and unspecified), dysthymic disorders, medical mood*

546 *disorders, current bereavement and depressive disorders not otherwise specified*