

## LETTER

doi:10.1017/S1041610215000769

**Comparing the Performance of the HADS and the GDS-15 in the AIBL Study**

Screening measures such as the 15-item Geriatric Depression Scale (GDS-15) (Sheikh and Yesavage, 1986) and the Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith, 1983) are important tools in the recognition of depressive symptoms in older people. While these measures are widely used, there is evidence of specific weaknesses in some cohorts and contexts, with the GDS-15 showing limitations in the context of cognitive impairment (Gilley and Wilson, 1997), and the depression subscale of the HADS (HADS-D) losing sensitivity in the context of older participants in hospital inpatient settings (Davies *et al.*, 1993).

We have conducted a study to compare the performance of the GDS-15 and the HADS-D in a large cohort of older people with varying cognitive profiles ranging from cognitive health to Alzheimer's disease (AD), and to investigate whether cognitive, health, or lifestyle factors demonstrate a relationship with GDS-15 and HADS-D scores. Analysis was performed on data ( $N = 972$ ) collected for the Australian Imaging, Biomarkers and Lifestyle (AIBL) study at its second time point in 2008–2009 (mean age 73.8 years; Range = 61.41–97.2) (Ellis *et al.*, 2009; 2014). The Healthy Control (HC) group comprised 692 participants; the Mild Cognitive Impairment (MCI) group 83 participants; and the AD group 197 participants, diagnosed by standard criteria as outlined in the original AIBL papers. Missing depression scores were concentrated among those participants with an AD diagnosis (25.9% GDS-15; 27.9% HADS-D), with fewer missing scores in the MCI group (10.8% GDS-15; 6% HADS-D) and HC group (4.2% GDS-15; 3.6% HADS-D). Additional variables analyzed were: Physical activity Metabolic Equivalent Task (MET)-minutes; the Mini-Mental State Exam (MMSE); and the Clinical Dementia Rating Scale sum of boxes score (CDR-SB).

Participants were grouped based on cut-off points for depressive symptoms, i.e. GDS-15 score of  $\geq 5$  and/or a HADS-D score of  $\geq 8$  (Bjelland *et al.*, 2002; Wancata *et al.*, 2006). Accordingly, participants were coded as “GDS+”, “GDS-”, “HADS+”, or “HADS-”. Participants were also assigned a binary code based on having

concordant or discordant HADS-D or GDS-15 scores. Concordant codes were assigned if participants had both GDS-15+ and HADS-D+ classifications, or alternatively GDS-15- and HADS-D- classifications. Discordant codes were assigned based on participants having HADS-D+ and GDS-15- classifications, or HADS-D-, and GDS-15+ classifications.

Forty-five participants (5.1% of the cohort) were identified as having depressive symptoms according to the HADS-D, while 82 (9.3%) were classified by the GDS-15 as having depressive symptoms. This relatively low prevalence likely reflects that participants had been screened to ensure they were not depressed before entry to the study at baseline. Not only did the GDS-15 identify the presence of depressive symptoms in nearly twice as many participants as the HADS-D, but it also identified these symptoms in a substantially different set of participants. Indeed, of the 71 GDS-15+ participants for whom a HADS-D score was recorded, 48 were classified as HADS-D-.

The measures' patterns of classification within each of the clinical groups (HC, MCI, and AD) were similar. According to both the HADS-D and the GDS-15, participants in the MCI and AD groups were more likely to be experiencing depressive symptoms than cognitively healthy older adults. The GDS-15 and HADS-D displayed similar patterns of correlation with other study variables. Age had a weak positive correlation with the GDS-15 ( $r_s(885) = 0.23$ ,  $p = <0.01$ ) and HADS-D ( $r_s(887) = 0.21$ ,  $p = <0.01$ ). MMSE score had a weak negative correlation with the GDS-15 ( $r_s(880) = -0.32$ ,  $p = <0.01$ ) and HADS-D ( $r_s(884) = -0.19$ ,  $p = <0.01$ ). The GDS-15 had a moderate positive correlation with CDR-SB score ( $r_s(880) = 0.41$ ,  $p = <0.01$ ), while the HADS-D had a weak positive correlation,  $r_s(884) = 0.22$ ,  $p = <0.01$ . The GDS-15 had a weak negative correlation with Physical Activity MET-Minutes ( $r_s(678) = -0.28$ ,  $p = <0.01$ ), while the HADS-D had a very weak to negligible negative correlation,  $r_s(686) = -0.17$ ,  $p = <0.01$ . Participants who received discordant GDS-15+/HADS-D- or GDS-15-/HADS-D+ classifications had similar demographic and cognitive profiles, according to the study's key variables. However, there were clear differences between the discordant and concordant participant groups. Those who received concordant HADS-D+ and GDS-15+ classifications tended to have lower MMSE scores, higher CDR-SB scores, and a lower level of physical

activity (reflecting the fact that they mostly came from the AD group). Logistic regression found that *better* cognitive function, as measured by higher MMSE or lower CDR-SB scores, was a significant predictor of measurement discordance. In other words, participants receiving discordant classifications appeared to be more physically active and more cognitively 'intact' than those participants who received concordant classifications.

An important finding is that, within the same dataset, the GDS-15 and HADS-D classified substantially different sets of participants with depressive symptoms. While the HADS-D's reduced number of cases may be explained by the narrower focus of its questions – primarily on anhedonia and depressed mood – this may not explain why the measures classified a substantially different set of participants. Indeed, both measures were *similarly correlated* with the age, cognitive, physical health, and physical activity variables. How, then, does this apparent equivalence sit with the fact that the two measures in fact captured a substantially different set of participants?

The difference between the measures' performance is most clearly illustrated by one of the study's most interesting findings. It was *better*, rather than impaired, cognitive function that significantly predicted measurement discordance. This finding helps explain the fact that discordance was more common in the HC than AD cohort, as the HC cohort was characterized by more intact cognitive function. A potential explanation for the unexpected direction of this finding is that people with cognitive impairment are likely to display more depressive symptoms than those who are not cognitively impaired, which might be a result of either depressive illness, or neurodegenerative disease that mimics or incorporates depressive symptoms. In contrast, those participants who are largely cognitively intact are likely to display less marked depressive symptoms, and occupy more of a 'middle ground' on the continuum from non-depressed to depressed. It appears that it is within this 'middle ground' of participants that the measures fail to agree.

With the HADS-D classifying fewer participants with depressive symptoms, this study may support previous research questioning the usefulness of the HADS-D in older cohorts. However, without reference to a gold standard test, we cannot rule out the possibility that the GDS-15 may be over-rating depressive symptoms. Further research examining more complex constellations of variables

may uncover the particular factors behind the measures' patterns of discordance.

## References

- Bjelland, I., Dahl, A. A., Haug, T. T. and Neckelmann, D.** (2002). The validity of the hospital and anxiety depression scale: an updated literature review. *Journal of Psychosomatic Research*, 52, 69–77.
- Davies, K. N., Burn, W. K. and McKenzie, F. R.** (1993). Evaluation of the hospital anxiety and depression scale as a screening instrument in geriatric medical inpatients. *International Journal of Geriatric Psychiatry*, 8, 165–169.
- Ellis, K. A. et al.** (2009). The Australian imaging, biomarkers and lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *International Psychogeriatrics*, 21, 672–687. doi:10.1017/S1041610209009405.
- Ellis, K. A. et al.** (2014). Rates of diagnostic transition and cognitive change at 18-month follow-up among 1112 participants in the Australian imaging, biomarkers and lifestyle (AIBL) flagship study of ageing. *International Psychogeriatrics*, 26, 543–554. doi:10.1017/S1041610213001956.
- Gilley, D. W. and Wilson, R. S.** (1997). Criterion-related validity of the geriatric depression scale in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, 19, 489–499.
- Sheikh, J. I. and Yesavage, J. A.** (1986). Geriatric dDepression scale (GDS). *Clinical Gerontologist*, 5, 165–173.
- Wancata, J., Alexandrowicz, R., Friedlich, F., Weiss, M. and Marquart, B.** (2006). The criterion validity of the geriatric depression scale: a systematic review. *Acta Psychiatrica Scandinavica*, 114, 398–410.
- Zigmond, A. S. and Snaith, P. P.** (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67, 361–370.

GUY CAMPBELL,<sup>1</sup> CHRISTINA BRYANT,<sup>1,2</sup>  
KATHRYN A. ELLIS,<sup>3,4,5</sup> RACHEL BUCKLEY<sup>1</sup>  
AND DAVID AMES<sup>3,5</sup>, THE AIBL  
RESEARCH GROUP\*

<sup>1</sup>School of Psychological Sciences, University of Melbourne, Melbourne, Australia

<sup>2</sup>The Royal Women's Hospital, Melbourne, Australia

<sup>3</sup>Academic Unit for Psychiatry of Old Age, Department of Psychiatry, The University of Melbourne, Australia

<sup>4</sup>Florey Institute of Neuroscience and Mental Health, Melbourne, Australia

<sup>5</sup>National Ageing Research Institute (NARI), Melbourne, Australia

\* Full list of the AIBL Research Group at [www.aibl.csiro.au](http://www.aibl.csiro.au).