The Depression Inventory Development Scale: Assessment of Psychometric Properties Using Classical and Modern Measurement Theory

A. Vaccarino1, A. Kalali2, S. Kennedy3, R. Lam4, R. Milev5, R. Uher6, P. Blier7, F. Farzan8, J. Foster9, A. Vaccarino1, A. Kalali2, S. Kennedy3, R. Lam4, R. Milev5, R. Uher6, P. Blier7, F. Farzan8, J. Foster9

Introduction

The Depression Inventory Development (DID) initiative1 aims to develop a comprehensive rating scale for MDD with more favourable psychometrically properties than current gold standards.2 Using an iterative process between field testing and psychometric analysis, an empirically-driven and collaborative protocol for item development was established. We report here results of latest iteration of DID items to assess symptoms related to anhedonia, cognition, fatigue, general malaise, motivation, anxiety, negative thinking, pain, sleep and appetite.

Methods

The present study was part of a larger multi-site, open label study conducted by the Canadian Biomarker Integration Network in Depression (CAN-BIND).2 Participants were administered 32 DID items using a semi-structured interview at each of two visits (MDD: Baseline, n = 211 and Week 8, n = 177; Healthy comparators: Baseline, n = 112 and Week 8, n = 104). Classical test theory, item response theory and Rasch measurement theory were applied to assess the psychometric properties of the DID items and determine which should be removed, modified or advanced. Participants were also administered the MADRS and QIDS-SR that allowed DID items to be evaluated against existing "benchmark" items. Data were managed on the Brain-CODE platform3

Results

Baseline demographic and clinical characteristics

MADRS (N=200) | QIDS-SR (N=200) | Did (N=211) | Did (N=177)
---|---|---|---
Female (%): | 132 (66) | 133 (63) | 177 (82)
Age in yrs, mean ± SD (range): | 45.74±10.72 (18-65) | 45.64±11.11 (18-75) | 43.47±10.78 (18-65)
MADRS ≥ 10 (range): | 298/85 (60–100) | 208/85 (60–100) | 298/85 (60–100)
QIDS-SR ≥ 10 (range): | 280/85 (60–100) | 208/85 (60–100) | 298/85 (60–100)

Classical Test Theory

The 32 DID items were examined for data quality using the following CTI acceptability criteria:

- Missing data <10%
- Endorsement of least/most variable option ≥40%
- Endorsement of any single option <50%
- Aggregate frequency endorsement of adjacent options ≥10%
- Low discrimination
- Inter item correlations/indiscernible >0.75

Based on CTI analyses, 8 items were removed from the item bank because of low variability in the range of response options, floor effects and/or skewness: Increased Appetite, Decreased Eating, Increased Eating, Headaches, Head-Bothered, Head-Intensity, Proroged Sleep and Weakness.

Item Response Theory

IRT was used to assess an item's ability (at the individual option level) to discriminate levels of depressive severity. Individual DID items were analysed separately against the ten MADRS items using Samejima's Graded Response Model4 in IRTPro Version 3.0. Option Characteristic Curves (OCCs) were generated to display the probability of selecting a particular option of a DID item as a function of overall depressive severity (MADRS). Discriminatory/slope (denoted as "a") provided quantitative measures of item discrimination, with slope parameters of a < 0.65 used as an indication of poor discrimination.

Rasch Measurement Theory

RMT (in WinSteps 4.2) was used to assess performance of the remaining items together as a scale. Person-item maps were generated that plot persons (individual participants) and items on a single continuum to compare the range and position of the person measure distribution to that of the item measure.

Additional DID Analyses

- Good internal consistency (α=0.95)
- Item-total correlations ranging from r = 0.49 (Decreased Appetite) to r = 0.84 (Drive)
- All items showed sensitivity to antidepressant treatment, with effects sizes ranging from r = 0.40 (Decreased Appetite) to r = 0.64 (Ahedonial)

Conclusions

C T I and IRT analyses were used to assess the performance of the DID items and identify which would be advanced, modified or removed. Fifteen DID items are recommended for inclusion in the penultimate scale, with modifications as required.

Acknowledgements

CAN-BIND is an Integrated Discovery Program carried out in partnership with, and financial support from, the Ontario Brain Institute, an independent non-profit corporation, funded partially by the Ontario government. Additional funding of the DID initiative provided by Industry. Previous iterations of DID were supported by the GSK Co-funding from Eli Lilly and Company. The authors declare that they have no competing interests.

References

2. Lam et al. BMC Psychiatry (2016) 16:105
4. Samejima. Psychometrika Monographs (1964) 39 (Suppl. 4)