ISCTM Fall 2018
Submission Date: 1 August 2018

Title: Evaluation of Relations between Negative Symptoms Scales in an IRT Framework

Authors: Anzalee Khan[1,2], Gregory P. Strauss[3], Daniel Umbricht[4], Philip D. Harvey[5], Danny Ulshen[2], Richard S.E. Keefe[2,6], Brian Kirkpatrick[7], Jean-Pierre Lindenmayer[1,8,9], William P. Horan[10,11] Anne M. Kring[12], Jack J. Blanchard[13], Racquel E. Gur[14], Michael F. Green[11,12], James M. Gold[15], Daniel N. Allen[16], Eric Granholm[17,18].

Affiliations: 1) Nathan S. Kline Institute for Psychiatric Research, 2) NeuroCog Trials, 3) University of Georgia, 4) F. Hoffmann-La Roche Ltd, 5) University of Miami, School of Medicine, 6) Duke University, School of Medicine, 7) University of Nevada, Reno School of Medicine, 8) New York University, School of Medicine, 9) Manhattan Psychiatric Center, 10) University of California, Los Angeles; 11) VA Greater Los Angeles Healthcare System, 12) University of California, Berkeley, 13) University of Maryland, College Park, 14) University of Pennsylvania, Perelman School of Medicine, 15) Maryland Psychiatric Research Center, 16) University of Nevada, Las Vegas, 17) University of California, San Diego, 18) VA San Diego Healthcare System.

Methodological Question:

The assessment of negative symptoms continues to be an important area of research in schizophrenia. Studies examining the psychometric properties of negative symptom scales have focused on estimates of reliability, validity, and factor analysis using classical test theory (CTT). These methods rely primarily on statistics that average across levels of individual variation and may obscure the fact that scale reliability is likely to vary across different levels of performance on the construct being measured. Examining the properties of individual scale items can lead to improvements in both the efficiency and reliability in mapping the full continuum of negative symptoms. Item Response Theory (IRT) methodology can help identify scales that work best across varying levels of severity and identify which items are most informative based on the severity of the study population. Identifying the correct clinical endpoint for a study population, including assessment of severity of impairment, can improve measurement precision and endpoint selection in clinical trials.

Aims:

We evaluated four negative symptom scales [CAINS, BNSS, Marder PANSS negative symptom factor (NSF), and the NSA-16] by: a) examining the performance of items at both the response (severity) and item (symptom) levels, b) examining the ability to discriminate individual differences in severity, and (c) addressing which scale provides the most information at different severity levels for each item.

Methods:

Data from screening or baseline evaluations were available for assessments in a sample of people with schizophrenia as follows: PANSS NSF, n = 7,186; NSA-16, n = 841; BNSS, n = 600, CAINS, n = 650. A non-

parametric IRT model with kernel smoothing approach was used to assess response and item characteristics for each scale.

Results:

For all scales, most items performed very well in terms of IRT criteria for item performance (100% of items on the BNSS, 83.33% of PANSS NSF, 68.75% of NSA-16, 84.62% of CAINS). The IRT suggested the following items performed poorly: Motor Retardation from the PANSS NSF; Reduced Daily Activity, Reduced and Inarticulate Speech, Reduced Social and Interest in Intimacy items of the NSA-16; and Motivation, and Expected Pleasure for Work and School on the CAINS. When compared to CTT methods, these items also showed poor item-total correlations. The Item Information Function showed that the BNSS, CAINS and NSA-16 were more informative (evaluates how individual test items measures across negative symptom severity levels) for subjects with moderate severity, compared to those with low or severe severity levels. In contrast, the PANSS NSF was equally informative for all levels of severity.

Conclusion:

This study provides additional evidence regarding the psychometric properties and clinical utility of four negative symptom measures, offers methodological contributions to the appropriate use of IRT, and helps elucidate variation in symptomatology. There were distinct advantages of each scale. The CAINS, BNSS, and NSA-16 were most informative for subjects with moderate severity levels, whereas the PANSS NSF was informative at all severity levels, although the BNSS, CAINS and NSA-16 have distinct conceptual advantages over the PANSS NSF and captures a broader spectrum of negative symptoms.