Reliability Findings from the Unified Global Training Program for the Brief Negative Symptoms Scale (BNSS)

Kirkpatrick B1, Opler M2-3, Strauss G4, Ahmed A5, Capodilupo G2, Tatsumi K2

1Dept. of Psychiatry & Behavioral Medicine, U. of Nevada-Reno, School of Medicine 2MedAvante-ProPhase, Inc. 3New York U. Medical Center 4Clinical Affective Neuroscience Laboratory, Georgia Psychiatric Risk Evaluation Program, U. of Georgia 5Psychotic Disorders Program. New York-Presbyterian/Westchester & Weill Cornell School of Medicine University Medical Center

BACKGROUND

- The Brief Negative Symptom Scale (BNSS) grew out of a recommendation from the NIMH-sponsored Consensus Development Conference on Negative Symptoms1. The BNSS consists of 13 items organized into six subscales2 and is rated based on a semi-structured interview with prompts and clarifying questions.
- The BNSS was designed to measure negative symptoms reliably and efficiently, and to be practical for use in clinical treatment trials. Key characteristics of the BNSS are that: (i) is brief, consisting of only 13 items that can be rated in a 10- to 15-minute interview; (ii) has a concise manual with a semi-structured interview guide; (iii) is written clearly and simply; and (iv) covers all five of the NIMH Consensus Conference domains; and (v) has good separation of the two dimensions thought to underlie negative symptoms (expressivity and anhedonia/amotivation/asociality).
- Reliability and other psychometric properties of psychiatric rating scales require careful evaluation across languages and cultures to prove their suitability for use in global clinical trials which require additional focus on the role of culture. This will help ensure that data can be properly interpreted across countries and regions to be pooled for analysis. The literature on this issue shows consistently that the expression of the same phenomena differs substantially across cultures4.
- Cross-cultural issues in assessment are particularly relevant in schizophrenia research5. Despite this, certain phenomenological structures are highly conserved, as seen in factor analyses of negative symptoms6. The BNSS has consistently demonstrated construct validity pertaining to avolition, anhedonia, and emotional expressivity7. Further, global studies of the BNSS have consistently demonstrated excellent internal consistency and validity with other negative symptom scales8, and recent research also reported that the BNSS is sensitive to drug effects, with effect sizes comparable to established scales8.
- The BNSS has been translated and tested in German, Italian, Russian, Polish, Japanese, Turkish, Spanish, French, Danish, Chinese, Norwegian, Dutch, Portuguese, and Korean9.

RESULTS

- In 2016, a global unified rater training program for the BNSS was launched across academic and industry users. Using pooled data, an analysis on the BNSS reliability was conducted on raters participating in multi-site trials (n=253).
- In this analysis (Figure 2) the BNSS demonstrated excellent reliability overall, with ICC of 0.903 [95 percent CI: 0.807-0.965, F(12,228)=10.35, p<0.001], with some variation among individual subscales.
- The Anhedonia subscale showed excellent inter-rater reliability with an ICC of 0.973 [95 percent CI: 0.884-0.999, F(2,26)=36.78, p<0.001]. When combined with item 4 for subjective distress, the ICC fell to 0.924 [95 percent CI: 0.741-0.959, F(3,57)=13.09, p<0.001].
- The Avolition subscale had a good ICC of 0.876 [95 percent CI: 0.688-1.000, F(19)=5.44, p=0.031]. Blunted affect was similar at an ICC of 0.843 [95 percent CI: 0.632-0.996, F(2,38)=6.377, p=0.004].
- The Asociality subscale also had a good inter-rater reliability with ICC of 0.893 [95 percent CI: 0.725-1.00, F(1,14)=9.33, p=0.009].
- One subscale, however, had a somewhat lower overall reliability, specifically the Asociality subscale with an ICC of 0.658, (F=2.923, p=0.104).

CONCLUSION / SUMMARY

- Raters from different countries may perceive and evaluate the symptomatology of schizophrenia differently due to differences in expression at the patient level, perception and interpretation among clinicians, and other contextual factors. To minimize the impact of cross-cultural effects, it is important to select and use a rating scale that reliably evaluates symptom severity and change across contexts.
- In this study, the BNSS achieved overall good-to-excellent inter-rater reliability among raters from different countries and cultures. The Asociality subscale showed comparatively lower reliability, suggesting there is a need for additional work to improve training, not only for the scale overall, but also for each subscale.
- The findings suggest that the BNSS is optimal for use in global clinical trials and further confirms that the BNSS is a valid and reliable negative symptom rating scale. Further study is warranted to investigate the cause of variability in Asociality, coupled with enhanced training and development of adjunctive strategies to improve inter-rater reliability across all subscales of the BNSS.