Rater Training on SANS and Monitoring of Rater Performance during Clinical Trials

Richa Gaur, PhD[1], Luis Ramirez, MD[2], Susan De Santi, PhD[3], Joep Schoemaker, MSc4

[1]The Cognition Group, Newark, DE, USA, [2]Quality Outcomes Training, Brecksville, Ohio, USA, [3]New York University School of Medicine, New York, NY, USA, 4Schering-Plough, Oss, Noord Brabant, Netherlands

Abstract

Introduction

Training a heterogeneous group of raters on the Scale for the Assessment of Negative Symptoms (SANS) is challenging as negative symptoms of schizophrenia are notoriously difficult to measure. This study evaluated ratings on the SANS and presents a novel methodology, Signal Enhancement System (SES), used in monitoring rater performance during the trial.

Methods

90 raters from 8 countries participated in a SANS rater training program via combining online and face to face training at the Investigators' Meeting (IM). Rater's performance was analyzed and inter-rater reliability and concordance with expert raters were measured. During the trial SANS patient interviews and ratings were monitored with the help of the SES methodology, which involved collaboration between the site rater (SR) and an independent rater (IR). Patient interviews were recorded at the site via a laptop and transmitted to a secure server and rated by an IR (blind to the site ratings) via the website. These independent ratings were shared with the SR to enable SANS score reconciliation before submitting to the sponsor.

Results

Inter-rater agreement for the total scores between expert and study raters ratings was substantial (kappa = .69, p < .001), moderate (kappa = .40 p < .001), in certification session I & II and poor (Kappa = .18, p = .06) in certification III, suggesting some raters experience difficulties administering the SANS. 95% raters achieved the passing grade (80% or higher). For the individual items the McNemar Test for paired binomial proportions showed raters agreement on only 8/25, 5/25 and 4/25 items with the expert ratings in certification sessions I,II and III respectively.

Conclusions

95% of raters certified on the SANS, but many raters were not able to achieve sufficient accuracy in ratings. Training alone does not ensure reliable ratings and statistical analysis of ratings from the site cannot establish whether the patient interviews were done appropriately in the study. To overcome the challenge of maintaining a high reliability of assessment during the trial, SES was implemented. This methodology helped in monitoring of recorded patient interviews which ensured compliance with interview techniques, provided an opportunity for SR performance feedback, and a forum for SR and IR calibration of scores.

Introduction

- Training was provided on Scale for the Assessment of Negative Symptoms (SANS, 25 items version) to investigators participating in multinational randomised controlled clinical trial.
- The SANS is challenging because negative symptoms are poorly understood and difficult to measure as they involve the absence of a symptom rather than the presence of one. Moreover, there is no standardized structured interview used in the administration of the SANS.
- We evaluated ratings using the SANS in a rater training program and present a novel methodology, Signal Enhancement System (SES), used in monitoring rater performance during the trial.

Methods

- 90 raters from 8 countries and 23 sites participated in a rater training program.
- The training program incorporated both online training via the web followed by live training at the Investigator Meeting (IM).
- A structured SANS interview was developed by the expert trainer and training provided to raters, raters had to be certified before being allowed to rate in the study.

Figure: 1 The Training and Certification Process

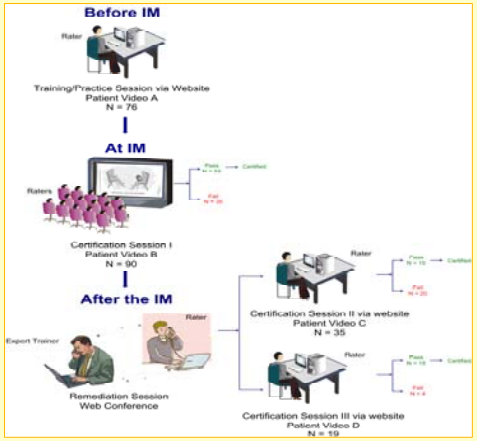


Table 1: Pass and fail percentage of raters in certification session I, II & III

Patient Video	N	Pass N (%)	Fail N (%)
B (certification session I)	90	55 (61)	35 (39)
C (certification session II)	35	15 (43)	20 (57)
D (certification session III)	19	15(79)	4 (21)

Rater agreement with the Gold Consensus Ratings (GCRs) on individual items

The McNemar test of Binomial Proportions;

- Certification session I: 8/25 items rated per the GCRs.
- Certification session II: 20/25 items achieved concordance with the GCRs.

Inter-rater reliability

- Inter-rater agreement was calculated using kappa coefficient, highest kappa level was observed in the certification session I (kappa = .69, p < .001; substantial agreement), and lowest in certification session III (Kappa = .18, p = .06 poor agreement).

Difficult to rate items

- Raters who failed the certification session had difficulty rating alogia which included poverty of speech, poverty of content of speech, blocking and increased latency of response.

Signal Enhancement System (SES)

- To maintain high reliability of assessment during the trial, SES was developed and implemented, with the following objectives;.
  1. Monitor patient interviews to ensure compliance with interview techniques.
  2. Provide an opportunity for Site Rater (SR) performance feedback and a forum for SR and Independent Rater (IR) calibration of ratings.
  3. Maintain SR's role at the key rater in the study.
  4. Controlling for rater's enrolment bias and expectancy bias

Who was involved in SES and how it works?

Site Rater (SR)

- Audio and video records the patient interview using a laptop with an inbuilt camera [proprietary software created by The Cognition Group (TCG)].
- Audio-visual recordings and collateral information transmitted to the TCG secure server.
- All patient ratings were recorded on a paper CRF and sent to a TCG secure central fax server.
- These scores were then uploaded to the secure web portal using ordinary internet access in the countries.
- Two IRs were selected from each country with extensive experience with rating scale and clinical trials.
- Recorded interview session received on the server was uploaded to the TCG website and the respective IR was invited to rate the video.
- The IR, blind to the ratings of the SR, reviewed the collateral information and provided his/her independent ratings by viewing the video online.
- Interviews at critical time points were recorded and reviewed for quality Rater Applied Performance Scale (RAPS) by the IR.
- SR invited to review the ratings of the IR and permitted to adjust the final score that is presented in the CRF document. SRs made the final rating decision, engaging the IR as a second opinion or "mirror".
- If the SR differed in the ratings outside set criteria, the SR had to provide a rationale as to the reason for not adjusting the site ratings.

The SES process is presented in figure 2, and SES portal is shown in Figure 3

Figure 2: Signal Enhancement System (SES) process flow chart

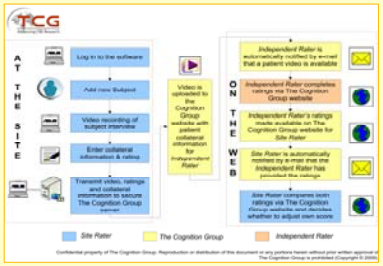


Figure 3: Signal Enhancement System (SES) Portal

Results of SES

- SANS baseline and endpoint comparison showed expectancy bias in clinical improvements (figure 5)

Figure 4: RAPS at baseline & endpoint

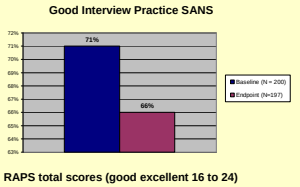
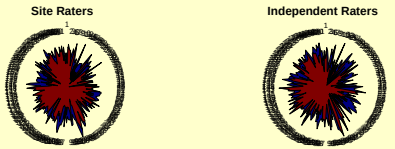


Figure 5: Clinical improvement based on SANS composite scores at Baseline (blue) and Endpoint (red) according to Site Raters and Independent Raters



Inter-rater reliability

- There is concordance in overall improvement of subjects between site raters and independent raters (figure 6).
- Despite the challenge of negative symptoms assessment, the results indicate that with an intensive training program and in-study monitoring (Hawthorne effect), sufficient inter-rater reliability can be attained.

Figure 6: Inter-rater reliability between Site Raters and Independent Raters Baseline (blue) and Endpoint (red)

	Intra-Class Correlation	Standard error of measurement
BASELINE N = 192		
$SANS_{1-22}$ - total score	**0.65**	**7.48**
Affective flattening (7)	**0.52**	**0.57**
Alogia (13)	**0.42**	**0.72**
Avolition-apathy (17)	**0.49**	**0.65**
Anhedonia-asociality (22)	**0.52**	**0.49**
Attention (25)	**0.40**	**0.83**
ENDPOINT N = 192		
$SANS_{1-22}$ - total score	**0.71**	**9.44**
Affective flattening (7)	**0.56**	**0.66**
Alogia (13)	**0.52**	**0.75**
Avolition-apathy (17)	**0.61**	**0.70**
Anhedonia-asociality (22)	**0.56**	**0.67**
Attention (25)	**0.62**	**0.66**

Conclusions

- Negative symptoms are a challenge to assess adequately with alogia being most difficult to rate.
- Training alone does not ensure reliable ratings and statistical analysis of ratings from the site cannot establish whether the patient interviews were done appropriately in the study or not. Signal Enhancement System was developed to monitor patient interview and ratings throughout the clinical trial.
- SES was well accepted by the ethics committees, patient, SRs, IRs and CRAs.
- SES was used in 8 countries some of which had rudimentary internet access, of 654 videos only 1.2% had technical problem leading to data loss. The technology is highly secure maintaining patient confidentiality.
- SES is an effective tool to control for bias, interview quality, inter- and intra-rater reliability (signal enhancement).
- The SES methodology with video recordings and second opinion ratings provides a new tool in the refinement and validation of rater scores; the principle "Investigator know their patients best".
- Adequate and appropriate training and monitoring using SES during the study leads to improved SRs performance.

References

1. Andreasen, N.C. (1984). Scale for Assessment of Negative Symptoms (SANS). Department of Psychiatry, University of Iowa. Iowa Ohio.
2. Lipsitz et al (2004). The rater applied performance scale: development and reliability. Psychiatry Research. 127:147-155
3. Kay, S.R., Opler, L.A., Fiszbein, A (1987). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophrenia Bull. 13: 261 -276
4. Schoemaker, J. H. (2008, November). *Signal Enhancement in Psychiatric Patients.* Presented at the Drug Information Association, Buenos Aires, Argentina.

The Cognition Group

254 Chapman Road, Topkis Building, Suite 100
Newark, DE 19702 USA
Tel: +1-302-454-1265, Fax: +1-302-454-7868
Toll Free: 888-454-1265

Presented at The International Society for CNS Clinical Trials and Methodology (ISCTM) - March 3-5, 2009 – Arlington, VA , USA