

RATER TRAINING: EVALUATION OF TWO DIFFERENT TRAINING METHODOLOGIES

Susan De Santi¹, Geetika Nath², Richa Gaur²

¹NYU School of Medicine, New York, NY, USA ; ²The Cognition Group, Newark, DE, USA

Abstract

Introduction

As primary outcome measures in CNS clinical trials rating scales significantly impact signal detection and study outcome. Rater training is incorporated into trials in an effort to increase the quality and reliability of clinical interviews. The design of the rater training programs may ultimately influence the quality of the clinical assessment. It is important to assess the effectiveness of programs, thus we report on two different methods for rater training.

Methods

Both methods included didactic training, opportunities to rate interview sessions with feedback, and certification sessions. Method 1, however was limited to e-learning (Website) with no live investigator meeting (IM), or remediation training. Method 2 incorporated e-learning and live training at the IM addressing difficult-to-rate items. One-to-one remediation was conducted for raters failing to certify. In both methods ratings were compared with expert gold consensus ratings (GCRs) and the pass/fail criteria were equivalent.

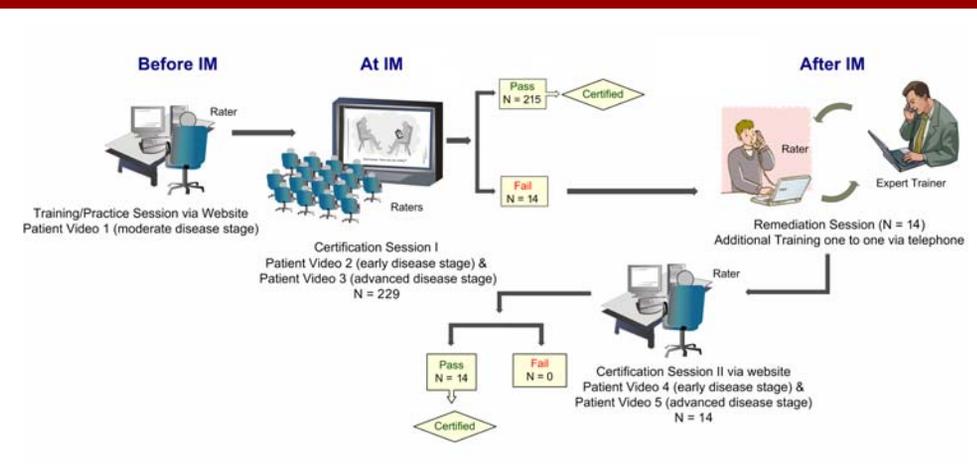
Results

In method 2, at the first certification opportunity, following discussion of difficult to rater items, 94% of raters certified, with 100% certification following remediation for those who failed. In method 1, 78% raters passed at the first certification (without focused discussion after the practice session). 8% of raters did not certify at all. Raters trained with method 2 showed greater inter rater reliability and agreement ($\kappa = .64 - .89$) with the GCRs, compared to method 1 ($\kappa = 0.28$).

Conclusions

Method 2 proved to be more effective. Targeted and focused learning activities lead to more effective ratings which may positively impact study outcome.

Figure 2. Method II Training and certification process



Introduction

- In clinical trials efficacy of treatment has traditionally been measured using rating scales designed to detect changes in symptom severity.
- The training and evaluation of clinical skills required for administering rating scales is essential before study raters should evaluate patients participating in a clinical trial.
- An aim of rater training is to increase rater agreement, reduce variance in raters ratings, and ensure uniformity in ratings scale administration across sites and raters.
- Factors associated with clinician ratings such as poor intra- and inter-rater reliability and interview quality, can significantly impact the study outcome.
- We present two different rater training methodologies performed in two separate trials, to assess the effectiveness of these training programs.

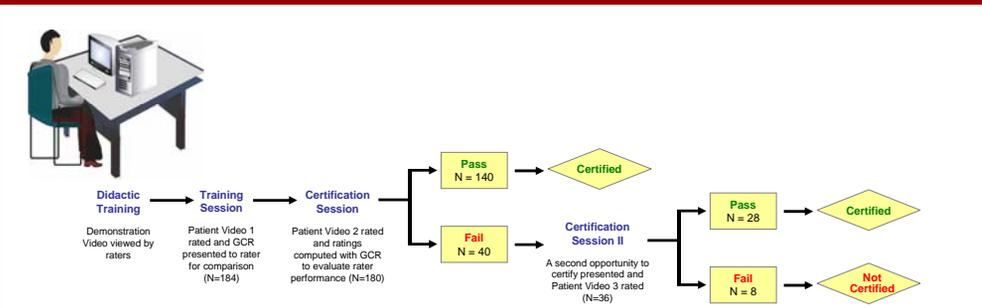
Methodology

Method I: Training was performed online via The Cognition Group's (TCG) secure training portal along with CD-ROMs (for sites with little or no connectivity). The training comprised three main sections:

- Didactic training using a demonstration video**
 - Scale background provided
 - Scale items presented
 - Scale administration and scoring demonstrated and guidelines for both provided
- Interactive online training/practice session:**
 - Videos of patients (played by actors) interviewed by an expert trainer were viewed by raters
 - Raters rated video and submitted ratings to TCG
 - Raters compared ratings with those provided by the experts in the field (known as gold consensus ratings (GCR))
- Certification:**
 - Raters rated new patient videos and submitted ratings to TCG
 - Raters' scores were compared to the GCR and performance evaluated to identify competent and inadequate raters
 - Raters who meet pass criteria are certified
 - Raters who failed, received another certification session

Method I training and certification process illustrated in Figure 1.

Figure 1. Method I Training and certification process



Method II: A global rater training program was designed which incorporated both online (web-based) training followed by training at a live Investigator Meeting (IM). Familiarization and training components were provided on all sections of the Rating Scale. Raters had to meet the pass criteria on two certification videos, early stage patient and advanced stage patient, to certify as a rater in the study.

Before the IM

- A study specific interactive training website was created which contained learning materials on the rating scale.
- Raters completed the first training session by observing a video of Patient 1 online and submitting their ratings online. Group ratings on Patient 1 were collected for analysis and discussion at the IM.

At the IM

- Training by an expert trainer was provided on all sections of the rating scale with emphasis on those items of variance which were identified during the online training program.
- The certification session at the IM included two videos: early stage patient and advanced stage patient.
- Raters had to meet the pass criteria for both the videos in order to certify on the scale. Raters viewed and independently rated the videos and entered ratings on rating sheets.

After the IM

- Raters' scores were compared to the GCRs and performance was categorized as pass or fail.
- Raters failing to certify for the early and advanced staged patient videos received individual remediation sessions with the expert trainer via the telephone and web conferencing facilities.
- After remediation, the raters were provided with another opportunity to certify with two additional videos of an early and advanced staged patient.

Method 2 training and certification process illustrated in Figure 2.

- For both methods the pass/fail criteria was provided by the experts based on an a priori predetermined level considered appropriate for a clinical trial.

Results

Method I:

- Pass Percentage of Raters:**
 - Certification Session I (N= 180): 78%
 - Certification Session II (N=36): 78%
- Rater Agreement on the Scale:**
 - Certification Session I (N= 180): Kappa indicated fair agreement (Kappa = 0.28 (z = 12.2, p<.001))

Method II:

- Pass Percentage of Raters:**
 - Certification Session I (N=229): 94%
 - Certification Session II (N=14): 100%
- Rater Agreement on the Scale:**
 - Certification Session I (N=229): Kappa ranged from substational to almost perfect agreement (Kappa = .64 - .89)

Conclusions

- Addressing difficult to rate items at the Investigator Meeting following the Training/Practice Session in Method II proved to be beneficial, leading to better performance of raters as compared to Method I.
 - 94% of raters certified in Method II versus only 78% in Method I.
 - Raters trained with Method II showed greater inter rater reliability and agreement ($\kappa = .64 - .89$) with the GCRs, compared to Method I ($\kappa = 0.28$).
- One-to-one remediation training proved to be beneficial as all raters who failed on their first attempt in Method II, passed on the second attempt following remediation sessions (via teleconference) focusing on specific items. In Method I some raters (12%) were unable to certify.
- Training Method II, a targeted, focused training program which evaluated rater performance after each session and addressed the reasons for poor performance after each session proved to be a more effective training method.
- Well structured and effective rater training programs can lead to more accurate ratings in a trial which may positively impact study outcome.

References

- Kobak, K.K., Lipchitz, J.D., Williams, J.B.W., Engelhart, N., and Bellew, K.M. (2005). A New Approach to Rater Training and Certification in a Multicenter Clinical Trial. *Journal of Clinical Psychopharmacology*; 25, 5: 407-11.

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