

Consensus Recommendations on Rater Training and Certification

by **THE CNS SUMMIT RATER TRAINING AND CERTIFICATION COMMITTEE: MARK D. WEST; DAVID G. DANIEL, MD; MARK OPLER, PhD, MPH; ALEXANDRIA WISE-RANKOVIC, PhD; and AMIR KALALI, MD**

Mr. West is Principal Consultant with Innovum Research, Las Vegas, Nevada (was with ePharmaSolutions, Plymouth Meeting, Pennsylvania, at the writing of this article); Dr. Daniel is Senior Vice President, Chief Medical Director, Bracket Global LLC and is with the Department of Psychiatry and Behavioral Sciences, George Washington University, McLean, Virginia; Dr. Opler is Chief Scientific Officer, Founder, ProPhase LLC, New York, New York; Dr. Wise-Rankovic is Vice President, CNS Clinical Development, INC Research, Raleigh, North Carolina; and Dr. Kalali is Professor of Psychiatry, University of California, San Diego, and Vice President, Global Head, Neuroscience Center of Excellence, Quintiles, San Diego, California.

Innov Clin Neurosci. 2014;11(11-12):10-13

ABSTRACT

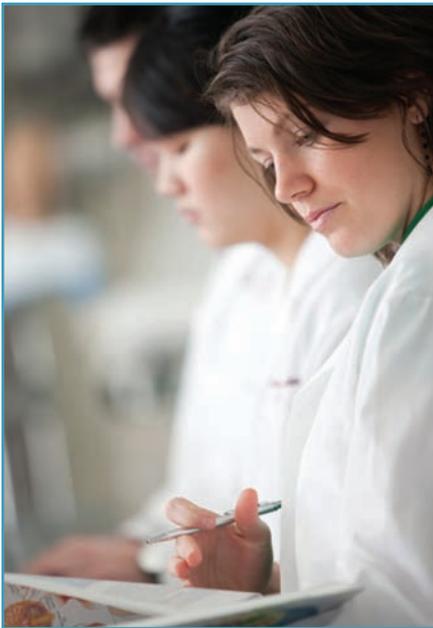
There is currently no accepted standard for the clinical research industry to follow when selecting and training raters to administer rating scales in clinical neuroscience trials. This article offers guidelines, based on expert recommendations of the CNS Summit Rater Training and Certification Committee, for selecting, training, and evaluating raters. The article also defines terminology and offers recommendations for considering raters with prior training and certification. These guidelines are intended for investigators, pharmaceutical companies, contract research organizations, and other entities involved in clinical neuroscience trials.

INTRODUCTION

Individuals with a wide range of skill and training routinely

administer rating scales used in central nervous system (CNS) clinical trials. Further, the training and certification methodologies used in clinical trials vary in their level of rigor. There is currently no accepted standard for the clinical research industry to follow when selecting and training raters to administer rating scales. Such scales are used as primary and secondary outcome measures contributing to the registration of investigational drugs and/or for empirical studies published in peer-reviewed journals. Studies such as these would be better served by an industry-wide guideline for rater training and certification.

The purpose of this manuscript is to define terminology (Table 1) and to propose general process guidelines for training, qualification, and certification of raters on clinician-rated scales commonly



FUNDING: No funding was provided for the preparation of this article.

FINANCIAL DISCLOSURES: Mr. West is a principal with Innovum Research, Las Vegas, Nevada. Dr. Daniel is Senior Vice President and Chief Medical Officer of Bracket Global LLC. Dr. Opler has received grant funding from National Institute of Mental Health (NIMH), National Alliance for Research on Schizophrenia and Depression (NARSAD), Stanley Medical Research Institute (SMRI), and Qatar National Research Foundation (QNRF); has received royalties from Multi-Health Systems; and is an employee of and shareholder in ProPhase LLC. Dr. Wise-Rankovic has no conflicts of interest relevant to the content of this article. Dr. Kalali has no conflicts of interest relevant to the content of this article.

ADDRESS CORRESPONDENCE TO: Mark D. West, 2441 Barclay St. Henderson NV 89044; Phone: (484) 557-5097; Email: mwest@innovumgroup.com

KEY WORDS: CNS, clinical research, rater training, certification, guidelines, recommendations, training methodology, rating scales

used in neuroscience clinical research. These guidelines are intended to be utilized as common framework for clinical trial investigators, pharmaceutical companies, contract research organizations, and other entities when establishing study protocol by providing a standardized baseline approach for selecting, training, and evaluating raters. The scope of the training recommended here is intended to serve a single project or protocol. Recommendations considering prior training and certification are included.

This article is a result of the work of the CNS Summit Rater Training and Certification Committee, which comprised experts in the field.

RECOMMENDATIONS FOR RATER QUALIFICATION

Determining minimum rater qualification. In general, sufficient qualification of raters comprises demonstrable experience with the administration of a particular rating scale or scales and a predefined term of relevant clinical interaction with the study subjects, as agreed upon by the key stakeholders involved in the study. The minimum qualifications for raters should be clearly documented prior to the start of the study. The study sponsor and/or the contract research organization (CRO) should consider these minimum qualifications during the site selection process as a reason for including, or excluding, a site or rater.

Training new raters. A research site may have raters who do not meet the minimum qualifications for one or more scales. In these cases, we recommend the following:

- Establish an agreement among the stakeholders on the training program to be developed for the project and ensure successful completion of the training program by raters before their approval.
- Complete in-field training.
 1. The principal investigator and/or a designated, qualified

sub-investigator who is certified to rate the scale in the study should document all mentoring of the raters in training. A standard for the type of documentation to be required should be established prior to the start of the study.

2. The principle investigator and/or a designated sub-investigator who is certified to rate the scale in the study should co-rate the study subjects with the rater(s) in training, until such time as the investigator certifies, in writing, that the rater is competent to administer the scale(s) on his or her own. The site should provide documentation of the co-rating process to the study sponsor to ensure proper compliance.

Note: The level of in-field training required for a study often depends on the amount of experience and training the rater has. Such training and mentoring

can range widely, and study-specific guidelines should allow for this range of prior experience so as not to prematurely exclude an appropriate rater simply on the basis of a guideline or requirement that is defined too narrowly.

RECOMMENDATIONS FOR PRIMARY STUDY MEASURES

This section contains our recommendations for training and testing raters prior to administration of scales that are of primary importance to the stakeholders. Scale importance may be based on its use as a primary outcome measure for regulatory submission or any other significant reason the study sponsor has for using the scale.

Minimum standards for training. Minimum training is dependent on early agreements about the training program with the relevant stakeholders. Rater training recommendations for each primary

TABLE 1. Rater training and certification terminology

Certification	The endorsements of an entity (e.g., business organization, sponsor, academic organization, or professional group) attesting to a rater's ability to properly administer a scale in the context of clinical research
Clinical neuroscience	An area of scientific research primarily served by the disciplines of psychiatry, neurology, pain, and neuropsychology
Qualification	The process of evaluating a rater against a defined set of requirements (i.e., qualification criteria)
Qualification criteria	The minimum educational, professional, and experiential credentials considered necessary for the competent administration of a scale by a rater.
Rater	A person involved with clinical research who administers a scale to a research participant
Scale	An instrument used to measure severity of signs and symptoms of disease or document diagnosis for purposes of clinical research; outcome measure
Testing	A process or examination involving a set of questions or example scenarios used as a means of evaluating the abilities, aptitudes, skills, and/or performance of an individual or group
Training	The educational process of establishing accurate, precise measurement of clinical trial endpoints among raters

study measure are as follows:

- Complete a didactic review of the purpose of the scale, standardized rules for administration, overview of some or many scale items, and the scoring for applicable items. A comprehensive review should be conducted for raters who are new to the scale being reviewed. This step may be waived for experienced raters who have demonstrated ability (such as through prior certification) depending on the rigor of the training program that is developed.
- Assess interview skills of the raters, including the following:
 1. Discussion of interview techniques;
 2. Demonstration of proper scale administration, if warranted, which may be done in person, using a recorded interview, or by other means using either a patient with or an actor trained to portray the disorder being studied.

Note: Patients' written consent should be obtained for use of any recorded interviews used in training.

Minimum standards for demonstrating competence.

Raters should demonstrate the following competency skills:

- Meet or exceed the minimum qualifications needed for the scale as defined by the study sponsor;
- Score one or more sample video interviews with a high degree of agreement with colleagues and/or expert consensus, with appropriate adjustments made for video quality and linguistic and cultural factors (degree of agreement and the basis for consensus, e.g., group or expert panel, may be determined *post hoc*).

A rater's established competence with scale administration, based on previous performance, may waive all other training requirements if agreeable to the stakeholders

involved. Such competence should be based on an established standard of testing and training that take into account appropriate skills and are in agreement with the consensus or expert panel. Provision for grandfathering may be considered should the rater have prior experience.

A comprehensive evaluation of a rater should include an assessment of proper administration of the rating instrument in a mock interview setting to establish the rater's ability to perform the skills learned during training. Sponsors who want to establish a rater's capabilities prior to study initiation may include this step.

GUIDELINES FOR OTHER SCALES

This section contains our recommended minimum standards for training and testing raters prior to administration of secondary outcomes, safety, and other scales as defined in the protocol for a clinical research study.

Training for raters should include a didactic review of the purpose of each scale, standardized rules for administration, overview of the scale items, and the scoring for each item. A comprehensive review should be conducted for raters who are new to the scale being reviewed. This step can be waived for experienced raters who have demonstrated ability (such as through prior certification).

CONSIDERATIONS FOR MULTINATIONAL STUDIES

Multinational studies introduce a number of variables that may affect study outcomes, and should be addressed as part of the training effort. The following variables may affect cultural validity:

- Linguistic differences in the scale version used in the study;
- Cultural and behavioral norms applied by clinicians;
- Clinical training and experience of raters with research trials.

Study sponsors should consider the implications of using a single

test case across all languages versus using multiple, culture- and/or language-specific examples for each culture and/or language group. A single video example provides a common basis for evaluating test data and necessitates some adjustment for each cultural and/or language group due to interpretational differences (e.g., translation and subtitling, culturally adjusted acceptable scores). Conversely, multiple culture- and/or language-specific examples allow for establishment of a higher degree of agreement within the culture and/or language group; however, this likely precludes any cross-cultural analysis. Study sponsors are encouraged to meet with their study statistician and make their rater evaluation process consistent with the statistical analysis plan of the study.

GUIDELINES FOR DOCUMENTATION

Training and certification should be documented for each site. In addition, a comprehensive training report should be prepared at the end of the study and maintained as part of study documentation.

Training methodology. The study sponsor or delegate should document the training methodology used for the study prior to the beginning of the study. For each scale, the document should specify the qualifications required, the contents of the training provided, and the testing methods used to determine certification (if required by the sponsor).

Site training records. Each site should receive a training record document that contains the following:

- Name(s) of rater(s) trained and/or certified;
- For each rater, the scale(s) trained and/or certified and date of training or certification;
- Sponsor name and protocol number;
- Name of trainer or training entity.

Site training records should be maintained at the site as part of the regulatory binder. This document should be reviewed by the study monitor on a periodic basis to ensure that ratings are being conducted by trained and/or certified raters (as specified by the sponsor, by scale). The study sponsor should also maintain a copy of each site's training record as part of their study documentation.

Raters may consider sharing their training records with sponsors in a shared database, such as the Global Rater Certification Database maintained by the CNS Summit.

Study documentation. At the conclusion of the study under the United States Food and Drug Administration (FDA) good clinical practice (GCP) guidelines,¹ a comprehensive report should be prepared for regulatory and scientific purposes. This report should document the qualification requirements for each scale used in the study, the training provided, and the certification results for all raters who participated in the study. The report should also document inter-rater reliability for scales for which testing data was collected, through statistical analysis such as Cohen's kappa coefficient, intra-class and/or

inter-class correlation coefficients (ICC), or Pearson's *r*.

RECOMMENDATIONS FOR RETRAINING AND RECERTIFICATION

We recommend that periodic retraining and/or recertification may be relevant to raters participating in longer-term studies.

Study sponsors should evaluate the burden placed on site personnel by retraining and/or recertification with the need to document ongoing inter-rater reliability during the study. Any requirement for retraining and/or recertification should be clearly stated to the sites prior to the beginning of the study.

Retraining, to include scoring conventions and guidelines, may be particularly desirable when a study is of longer duration and the frequency of scale administrations is low.

Study sponsors might wish to consider the implications of recertifying raters during a study. For example, if a rater no longer meets certification criteria, what does this mean for the study data previously collected by the rater? Is the rater allowed to continue in the study? Study sponsors should make the decision regarding their

approach prior to beginning the study.

All retraining and results of recertification activities should be documented by the study sponsor and maintained as part of the study documentation.

SUMMARY

Obtaining valid, reliable, and accurate ratings of patient symptoms in CNS clinical trials is of vital importance to study success. Common standards for rater training and certification have not previously been established. This manuscript provides recommendations to establish a common minimum standard across CNS studies.

REFERENCE

1. United States Food and Drug Administration. Science & Research. Regulations. FDA regulations governing the conduct of clinical trials describe good clinical practices (GCPs) for studies with both human and non-human animal subjects. <http://www.fda.gov/ScienceResearch/SpecialTopics/RunningClinicalTrials/ucm155713.htm>. Accessed December 19, 2014. ■