

The impact of rater training on clinical outcomes assessment data: A conceptual review

Sadler ME¹, Yamamoto RT¹, Khurana L¹, Dallabrida SM¹

¹eResearch Technology

Methodological Question Addressed

This review aimed to evaluate the empirical evidence for the impact of rater training on the quality of clinical outcomes assessments (COAs).

Introduction (Aims)

COAs are essential to data collection in clinical trials across a wide range of indications, and the quality of such data may determine the success of a trial. However, the methods for ensuring accurate and reliable assessment in clinical trials are often unreported. Training of those individuals who complete COAs, including site personnel, patients, and caregivers is recommended by experts in the field, and standards have been proposed for training, evaluation and monitoring of rater performance. Regulatory authorities increasingly seek evidence of training in the evaluation of labelling claims based on COA data.

The purpose of this review was to determine the extent and nature of the empirical evidence for the benefits of rater training on the quality of COA data.

Methods

A targeted review of 1,891 articles published 1993 - 2016 identified 29 eligible papers containing data on rater training and the reliability and accuracy of COA data applicable to clinical trials in any therapeutic area. Studies reported training of site staff; only 1 study reported effects of patient training. Rater training elements were classified as didactic (video or live lecture on administration and scoring), practical (practice scoring of interviews or other stimuli to a “gold standard”), and applied (conducting and scoring an interview with live or remote observation and feedback, with remediation if needed).

Results

The majority of studies were in psychiatry (n=14) for depression using the Hamilton Depression Scale and schizophrenia using the Positive and Negative Symptom Scales, followed by neurology (n=5), psoriasis (n=3), drug-induced movement disorders (n=2), and miscellaneous (n=5). All studies reported some form of didactic instruction, most accompanied by practical training (n=18) and/or applied training (n=4).

Of the studies that included a comparison group 13 showed improvement while 4 found no differences with training. In the 11 studies lacking a comparison group it was not possible to determine whether training improved rater skills or data quality due to a lack of relevant data.

Conclusions

The findings indicate that rater training is associated with significant improvement in the accuracy and reliability of COAs across diverse indications and instruments when training meets certain standards. The following conclusions are supported:

- Rater training is effective in improving data quality as measured by inter- and intra-rater reliability and accuracy regardless of raters' discipline, education level, credentials or level of previous experience.
- Training should provide clear anchor points and objective scoring criteria across the range of possible scores on each item of an instrument.
- 76% of studies that included a comparison group, demonstrated improved inter- and intra-rater reliability following any combination of training components (didactic, practical and applied).

Findings are discussed in the context of escalating clinical trial duration, complexity and cost as well as regulatory guidance for labelling claims. Recommendations for further more rigorous studies are discussed.

Disclosures

One or more authors report potential conflicts of interest which are described in the program.

The authors are employees of ERT.