Data Quality Metrics for Clinical Outcome Assessments in Alzheimer's Disease Clinical Trials

E. Bartolic[1], A Schembri[2], C. J. Edgar[3],

[1]Cogstate, Clinical Science, New Haven, USA; [2]Cogstate, Clinical Science, Melbourne, Australia; [3]Cogstate, Clinical Science, London, UK

Background

Data quality is recognized as a significant concern for clinical outcome assessments (COAs) in Alzheimer's disease (AD) trials. Complexity and lack of standardization has led to numerous data quality initiatives and recognition of the value of objective performance-based outcome (PerfO) assessments and biomarkers. However, clinician reported outcome (ClinRO) COAs remain prevalent and strategies to identify and reduce errors critical to trial conduct. Per FDA guidance, risk-based approaches to central monitoring are encouraged to focus effort on the most important aspects of data quality via risk assessment, monitoring plan development, corrective action plan and central monitoring.

Development and selection of reliable COAs is important to trial conduct since reliability and its impact on variance are a determinant of trial sample size. Such properties are dataset dependent not invariant and efforts to standardize training and data quality monitoring are also critical.

This is even more important for ClinRO and PerfO assessments, given their complex administration and scoring needs and at least one recent trial using the CDR-SB (ClinRO) as a primary outcome increased sample size due to unexpectedly high variance.

Methods

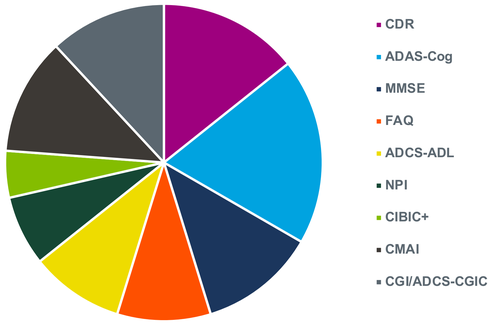
The aim of this analysis was to review COAs used in the past 5 years in Phase 3, AD clinical trials, and investigate the most commonly used, data quality initiatives adopted, and reported data quality metrics.

Clinicaltrials.gov was searched to identify COAs used in industry sponsored, Phase 3, clinical trials in AD over the past 5 years (N=49 trials identified).

A search of pubmed over the same period identified publications related to data quality/error reduction strategies and analyses of reliability data for clinical trial COAs. A further search of AAIC abstracts over the same period also identified publications related to data quality/error reduction strategies.

Clinical Outcome Assessments

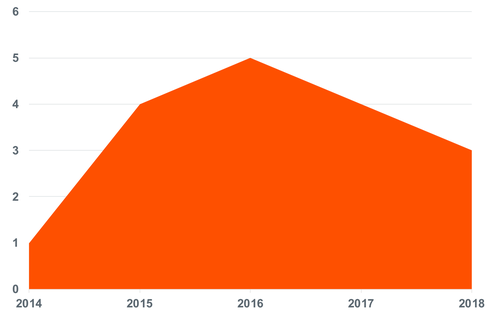
Figure 1: Most commonly used COAs



- Alzheimer's disease clinical trials in recruitment at the time of the clinicaltrials.gov search (N=18) reported use of 13 different COAs, nine of which were used in two or more trials.
- The most commonly used (ADAS-Cog, 8 trials; and CDR, 6 trials) are a mix of clinician-reported (ClinRO) and performance-based (PerfO) outcome assessment, with complex administration and scoring requirements.
- Such instruments present reliability issues. Although CDR-Global Score reliability has been reported as high, studies have identified low inter-rater reliability under certain circumstances that might particularly impact CDR-SB e.g.:
  - Patients at the questionable and mild dementia stages
  - Domains including orientation, judgment, home and hobbies, and personal care
  - Naïve raters
    (Rockwood et al 2000; Tractenberg et al, 2001)

Reporting of data quality initiatives

Figure 2: AAIC posters reporting data quality initiatives by year



- Data quality initiatives are a regular feature at AAIC poster sessions, see Figure 2.
- However, these invariably report error rates, and none report reliability or variance metrics.
- Given that error rates are dependent on the data quality initiatives employed (e.g., training and qualification; worksheet reviews alone vs. those including audio and video reviews; paper source vs. eCOA), the error metric cannot be compared across different trials.
- Furthermore, the success of these initiatives may in part be determined by their impact on reliability and variance, which is rarely reported.
- Very little is currently known about the impact of data quality initiatives on outcome measure reliability and variance.

Conclusions

- Reliability and variance remain important areas of concern for trial conduct. Reported psychometric properties are important in the selection of scales and data quality initiatives provide critical information about trial conduct.
- The association between error rates, specific data quality initiatives and the reliability and variance of different outcome measures is not well understood.
- A focus on increasingly comprehensive initiatives to improve detection of errors has an uncertain impact on reliability/variance and is not well aligned with risk-based monitoring.
- Although statistical modeling may provide one route towards a more risk-based approach to data quality, error rates in statistically aberrant data may not be high (e.g., of 12% [N=750] of statistically aberrant data, only 20% [N=152] of these had at least one identified scoring or administration error) - *Karas et al, 2018*
- Routine reporting of reliability and variance metrics is an important part of understanding success of trial conduct, including data quality initiatives.
  - This is aligned with good practice recommendations e.g.,
    - Reporting of inter-rater reliability (West et al, 2014 THE CNS SUMMIT RATER TRAINING AND CERTIFICATION COMMITTEE)
    - Reporting of standard deviation (CONSORT 2010)
- Such metrics can be explored and reported during trial conduct using screening and baseline data and comparison of raters against 'gold-standard' assessment conduct, as well as post trial completion and test retest reliability is a further key metric that should be routinely reported.

Alzheimer's Association International Conference; July 2019; Los Angeles, USA

Scan to download a file of this poster