SIGNANT HEALTH
THE EVIDENCE GENERATION COMPANY™

The Signal

Selecting and Monitoring Clinically Meaningful Data Quality Indicators

The ongoing, accelerated digital transformation of clinical trials presents new opportunities for medical research sponsors and clinical research organizations. Increasing adoption of electronic clinical outcome assessment (eCOA) and other data capture technologies, coupled with new approaches in study design, help researchers collect more data on more endpoints than ever before. By applying advanced analytics tools and methodologies, these data can be analyzed to explore clinically meaningful concepts and reveal specific insights.

At Signant, we focus on early identification of data concerns to be able to address causes ideally in the screening period. Our Blinded Data Analytics solution helps study teams statistically monitor and analyze clinician-defined quality indicators such as anomalies, unusual patterns, or statistical outliers that can represent data of questionable credibility. We begin by defining the quality indicators for each study. Next, data are aggregated from different sources such as eCOA systems, paper-based assessments, and/or audio recordings. By comparing these study data to other relevant data sets at the rater, site, study, country, and region level, we can reveal potential sources of low-quality data. Study teams can investigate the data quality indicators further and intervene to correct problems, improving the quality and integrity of endpoint data.

The quality indicators chosen for evaluation throughout the course of a study depends on several factors such as the indication, scales or instruments used, protocol requirements, and sponsor preferences. Below are some examples of quality indicators and their implications in terms of potential impact on endpoint reliability.

Anomaly Detection

Leveraging machine learning algorithms, we can review many characteristics at site and participant levels to identify patterns of concern. For example, an analysis of patient diary PIN codes can reveal potential fraud. If diary PIN codes repeat, it could suggest that a site is influencing the selection of PIN codes by suggesting, for instance, that participants use their birth year as a PIN code. While seemingly harmless, in theory the site staff can deduce PIN codes and either enter or alter data. Matching or sequential PIN codes could also indicate that a site has falsified participants.

Erratic Ratings

Large changes in scale scores over the course of two or more visits could suggest that raters

need more training, or it could reflect symptom instability in a subgroup of participants. In one recent example, we evaluated erratic ratings in a schizophrenia trial and compared it to 22 similar schizophrenia trials, which helped our study team determine that expectation bias, increased placebo response, and measurement error contributed to the unexpected variations.

Score Discordance

An analysis of instrument scores from visit to visit can reveal discordances in point change values that may suggest administration or scoring errors by raters. When Signant's data analytics team found clinically meaningful discordance in over 11% of study visits in one Alzheimer's trial, they expected to find scoring errors. However, because they were also able to evaluate data from independent endpoint quality reviews of the assessment worksheets and audio recordings, it was clear that there was no correlation between the score discordances and instrument administration or scoring errors. This means they can look elsewhere for other causes of the discordance.

These are just three of many potential quality indicators that can be monitored to improve the reliability of endpoint data a study generates. Others include score outliers or inliers, rater change, interview duration, variance checks, missing visits, and identical scoring, to name a few. If present, quality indicators usually reveal one of several common threats to data quality: the need to improve rater consistency and accuracy through targeted training, fraud, or study design or protocol elements impacting data quality such as eligibility criteria.

If left unaddressed, these quality indicators can increase noise, making signal detection more difficult, and potentially result in negative or failed trials. A well-honed data quality management program consisting of statistical monitoring, independent reviews, and ratings quality management has been proven to increase data reliability.

Contact our data management team to discuss endpoint reliability solutions for your program or protocol, and subscribe to our blog to stay apprised of important data analytics topics and trends.

Alan Kott, MUDr.

Clinical Vice President and Practice Lead, Data Analytics