# Data Quality Monitoring of Bayley Assessments: Error Detection in Rare Neurodevelopmental Disorder Trials

**Submitter** Laura Miller

**Affiliation** Cogstate

## SUBMISSION DETAILS

**I agree to provide poster pdf for attendee download.** Yes

**I have used the poster abstract template to develop my abstract.** Yes

**Methodological Issue Being Addressed** Rare neurodevelopmental disorder clinical trials rely on complex developmental outcome measures, where even minor administration or scoring errors may disproportionately affect data interpretability due to small sample sizes. Bayley-III assessments are particularly vulnerable to such errors given their procedural and scoring complexity. This study evaluates a structured data quality monitoring framework designed to systematically identify, classify, and correct Bayley administration and scoring errors while ensuring reviewer reliability through standardized training and calibration procedures.

**Introduction** Neurodevelopmental outcome measures such as the Bayley-III require high inter-rater reliability and strict adherence to standardized administration procedures. In trials of rare disorders, decentralized data collection across sites with variable assessor experience introduces risk for administration and scoring errors. While centralized data quality monitoring approaches are increasingly adopted, relatively few studies describe how identified errors are evaluated, corrected, and quality-controlled at the reviewer level. This analysis examines error frequency, magnitude of score change, and reviewer reliability across Bayley-III administrations drawn from two rare disease clinical trials.

**Methods** Data were aggregated from two rare neurodevelopmental disorder clinical trials. Each administration underwent structured data quality review by trained clinical reviewers. Reviewers completed standardized training and annual recalibration against gold-standard reviews, with all reviewers meeting the predefined performance threshold of >90% agreement.

Descriptive analyses were performed across administrations. The primary endpoint was the proportion of administrations requiring score correction following data quality review.

**Results** Across 133 total administrations (Trial A: n = 54; Trial B: n = 79), 51.13% (n = 68) contained no identified errors, while 48.87% (n = 65) had at least one error. Data quality monitoring identified 29 administrations requiring score correction, representing 21.80% of all administrations and 44.62% of administrations with at least one identified error.

Score correction magnitude was evaluated to assess potential impact. One corrected administration reflected an identified error in which a missing score was assigned a value of zero and was excluded from magnitude categorization due to the absence of a numerical score change. Among

the remaining 28 numerical corrections, 22 (78.57%) were categorized as small (1–2-point change in raw score), 4 (14.28%) as moderate (3–5-point change), and 2 (7.14%) as large (>5-point change).

**Conclusion** This structured data quality monitoring framework effectively identified and resolved Bayley administration and scoring errors across rare neurodevelopmental disorder clinical trials. Quantification of correction magnitude and formal reviewer calibration strengthened methodological rigor and data interpretability. In rare neurodevelopmental disorders, even small score changes may be clinically meaningful; therefore, it is critical that assessments capture the most accurate representation of a participant's current level of functioning. These findings support integrating structured data quality monitoring with reviewer reliability controls as a best practice for trials employing complex developmental outcome measures.

Limitations: This evaluation relied on descriptive analyses and did not assess downstream clinical significance of corrected scores, differences across individual trials or sites, or associations with rater experience. Future analyses will examine these factors and the impact of corrections on derived outcomes.

### Co-Authors

**Laura Miller[1],** Whitney Wood[1], Pam Ventola[2]

[1] Cogstate

[2] Cogstate; Yale Child Study Center

### Keywords

| Keywords |
| --- |
| Data Quality Monitoring |
| Rater Reliability |
| Outcome Measure Validation |

**Guidelines** I have read and understand the Poster Guidelines

**Disclosures** None.