From Paper to Precision: Improved MADRS Scoring Accuracy with eCOA Implementation

Authors **Rumerman, S, Brady, C, Negash, S, Seagal, I, Opler, M**

Clario, eCOA Clinical Science and Consulting, Philadelphia, PA, USA

Introduction

Clinical trials in depressive disorders often suffer from high failure rates, partly due to variability and noise in endpoint data. A significant contributor to this variability is rater error, which undermines interrater reliability and internal consistency. Traditional paper-based clinical outcome assessments (COAs) are particularly vulnerable to such errors, requiring manual calculations, transcription into electronic data capture (EDC) systems, and extensive source data verification. These processes increase administrative burden and introduce multiple opportunities for error.

Electronic COA (eCOA) platforms offer several advantages, including automated scoring, real-time data validation, and the elimination of transcription errors. The tablet-based eCOA system utilized captures source data digitally, includes consistency checks assembled from the International Society for CNS Clinical Trials and Methodology (ISCTM) working-group[1] and provides real-time clinical guidance to standardize scale administration and scoring, thereby improving data quality. This study aimed to compare scoring error rates between paper-based and eCOA-based administrations of the Montgomery–Åsberg Depression Rating Scale (MADRS), a widely used clinician-rated scale for assessing depression severity[2].

Methods

MADRS assessments from two multinational Phase II/III depression clinical trials were analyzed: one using traditional paper-based administration (n = 1,100) and the other using the eCOA platform (n = 1,100). In both trials, site raters completed the interviews using the Structured Interview Guide for the MADRS (SIGMA) and assessments were centrally reviewed by expert reviewers using audio recordings and source documentation. Discrepancies between recorded scores and expected scores, based on interview content and standardized scoring conventions were identified.

The percentage of assessments with at least one discrepancy, and those with two or more discrepancies, were compared across modalities. Item-level discrepancies between eCOA and paper modalities will be examined, with summary statistics presented to include the number and percentage of discrepancies by type. The analysis identifies the most common errors in each modality and evaluates whether eCOA mitigated specific types of errors more effectively than paper.

Results

Percentages of reviews with at least one scoring discrepancy, as well as those with two or more discrepancies, were substantially lower in eCOA administrations compared to paper-based (97% to 30% for at least one discrepancy, and 90% to 13% for two or more discrepancies). Analysis of variance (ANOVA) comparing the number of errors per administration by scale revealed significantly lower error rates on eCOA versus paper-based administration, $F(1, 2198) = 2080.90$, $p < .0001$.

- Figure 1 shows the percentages of reviews with discrepancies for paper-based and eCOA administrations. Percentages of reviews with at least one discrepancy, as well as those with two or more discrepancies, were substantially lower in eCOA administrations compared to paper-based

- Analysis of variance (ANOVA) comparing the number of errors per administration by scale revealed significantly lower error rates on eCOA versus paper-based administrations.

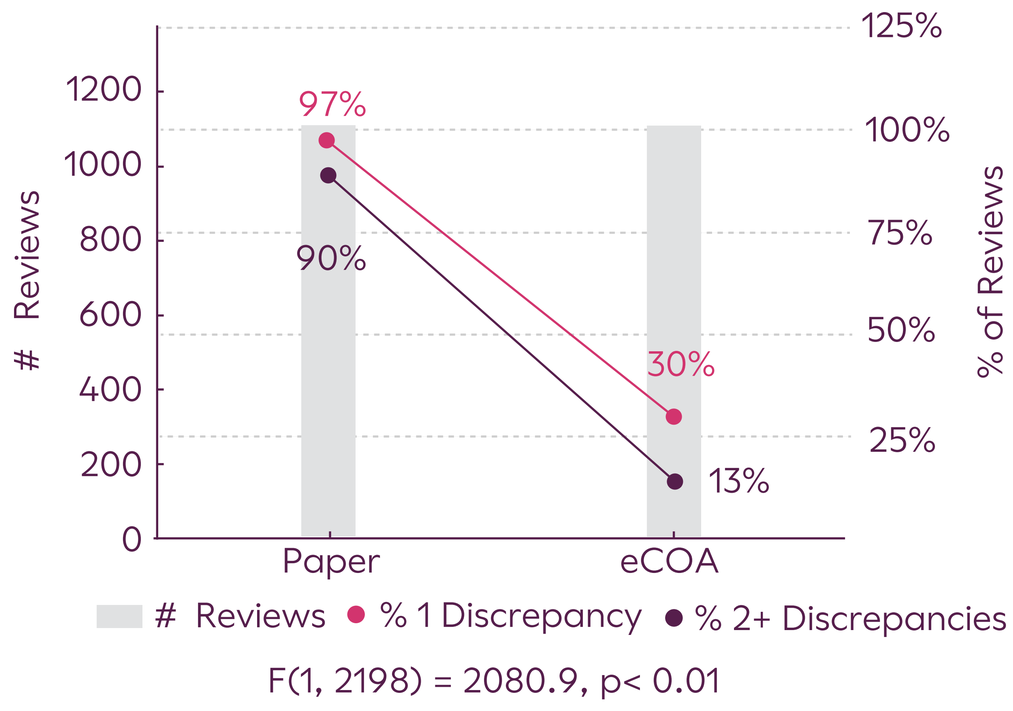

Figure 1: % Reviews with Discrepancies by Administration Type

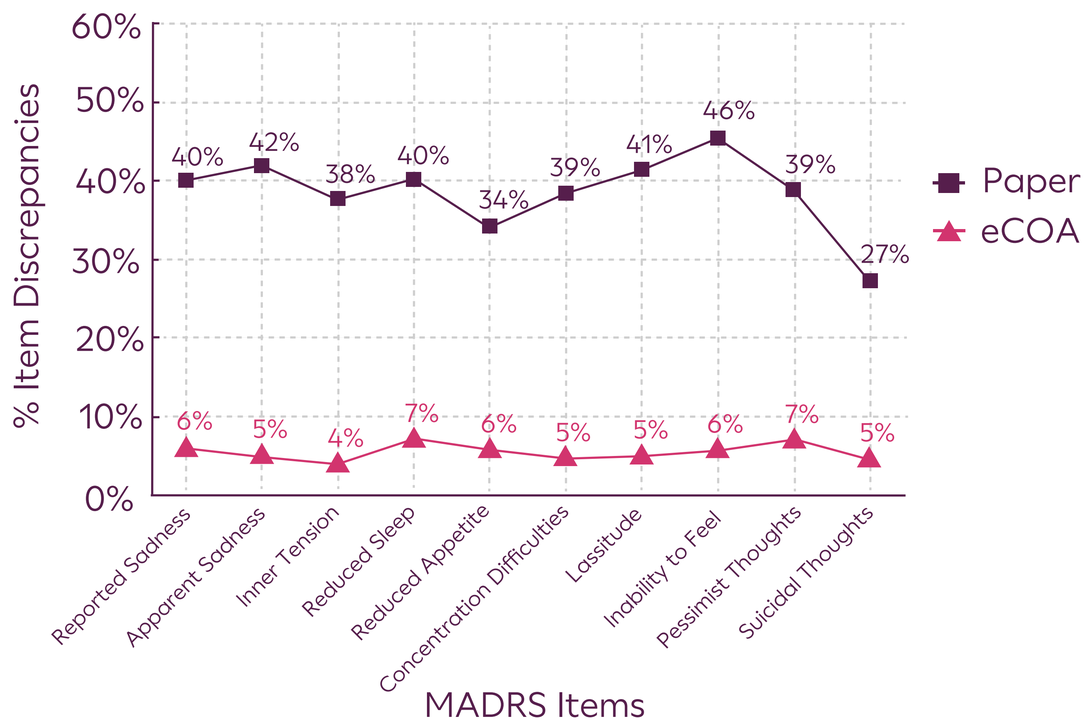

Figure 2: % Items with Discrepancies by Administration Type

- Figure 2 shows item-level score discrepancies. eCOA administrations resulted in consistently lower error rates than paper-based assessments for all items.

- eCOA significantly reduced scoring discrepancies in all items. For example, scoring discrepancies for “Inability to Feel” dropped from 46% in paper-based assessments to 6% with eCOA, and “Apparent Sadness” discrepancies decreased from 42% to 5%.

- “Inability to Feel” should be rated relative to the patient’s Euthymic Baseline, which is collected during the initial administration in both paper and eCOA formats. However, eCOA provides raters with a reminder of the patient’s reported euthymic baseline timeframe for reference, whereas paper-based assessments do not include this feature. This feature, along with the eCOA pop-up guidance for administration and scoring, may help to improve scoring consistency.

Conclusion

The use of eCOA significantly reduces scoring inconsistencies in MADRS assessments compared to paper-based methods. These findings support the clinical and operational advantages of eCOA platforms, including automated error detection, reduced site burden, and improved data quality. While methodological differences such as rater training and study populations may have influenced results, the substantial reduction in error rates - ranging from 69% to 86% - demonstrates the value of eCOA in enhancing the reliability of psychiatric clinical trials. These features collectively contribute to more accurate, efficient, and scalable clinical trial operations.

References

1.Rabinowitz, J., Schooler, N. R., Brown, B., et al (2019). Consistency checks to improve measurement with the Montgomery-Asberg Depression Rating Scale (MADRS). Journal of Affective Disorders, 256, 143-147.

2.Montgomery, S.A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. British Journal of Psychiatry, 134(4), 382–389.

CLARIO.