

Rater Applied Experience during Start-up Predicts Performance during Alzheimer’s Disease Clinical Trials

Presenters: Katrina Patrick, PhD, Bryan Karazsia, PhD, Cristian Sirbu, PhD, Jennifer Olt, PhD, Cynthia McNamara, PhD
All Authors/Presenters are affiliated with IQVIA Rater Services, formerly Cronos Clinical Consulting Services, Inc.

Introduction

Site raters who administer and score Clinical Outcome Assessments (COAs) are often qualified for trial ratings via prespecified thresholds for their education and prior experiences. Following qualification, they undergo trial and COA-specific training for synchronization of administration and scoring. Limitations of relying on educational and experience thresholds are documented (e.g., Kobak et al., 2004), and emerging evidence suggests that rater performance on applied training exercises is a more proximal indicator of in-trial rating quality (e.g., Patrick et al., 2025).

The present study examined whether rater qualifications and performance on a standardized quiz or applied scoring exercise during trial startup predict in-trial rating performance in Alzheimer’s Disease (AD) clinical trials.

Methods

Data were synthesized from 342 raters across four AD clinical trials. Because some raters contributed ratings for multiple outcome measures and/or across trials, 477 observations were available for analysis. We accounted for this non-independence by using mixed-effects logistic regression to model in-trial performance (whether or not raters required in-trial remediation due to administration or scoring errors). Predictors included AD clinical experience (years), education level, scale-specific experience (years and recent administrations), and the number of attempts required to pass the applied scoring exercise.

Conclusions

These results show that many of the assumptions made during rater selection may need to be reexamined. Basic educational level is not predictive of rater performance in AD studies, calling into question that higher degrees lead to higher quality. The number of recent scale administrations is also a poor predictor. In fact, the range of administrations reported by raters is so wide, the standard deviation is larger than the mean. Clinical experience with the population is also not predictive. Finally, experience with the scale is predictive of rater performance with more experience resulting in a slightly lower odds of requiring in-trial remediation.

- By far the best predictor of in-trial rater performance is their performance on the training task, either a quiz or a scoring exercise. Perhaps in the future, applied performance tasks could act as a rater selection tool as well as a training exercise.

Disclosures: All authors report no conflicts of interest for this work. All are current employees of IQVIA Rater Services, formerly Cronos Clinical Consulting Services, Inc. IQVIA does provide Rater training, Rater oversight and Rater intervention services.

Alzheimer’s Association International Conference – July 12-15, 2026 – London, UK

© 2026. All rights reserved. IQVIA® is a registered trademark of IQVIA Inc. in the United States, the European Union, and various other countries.

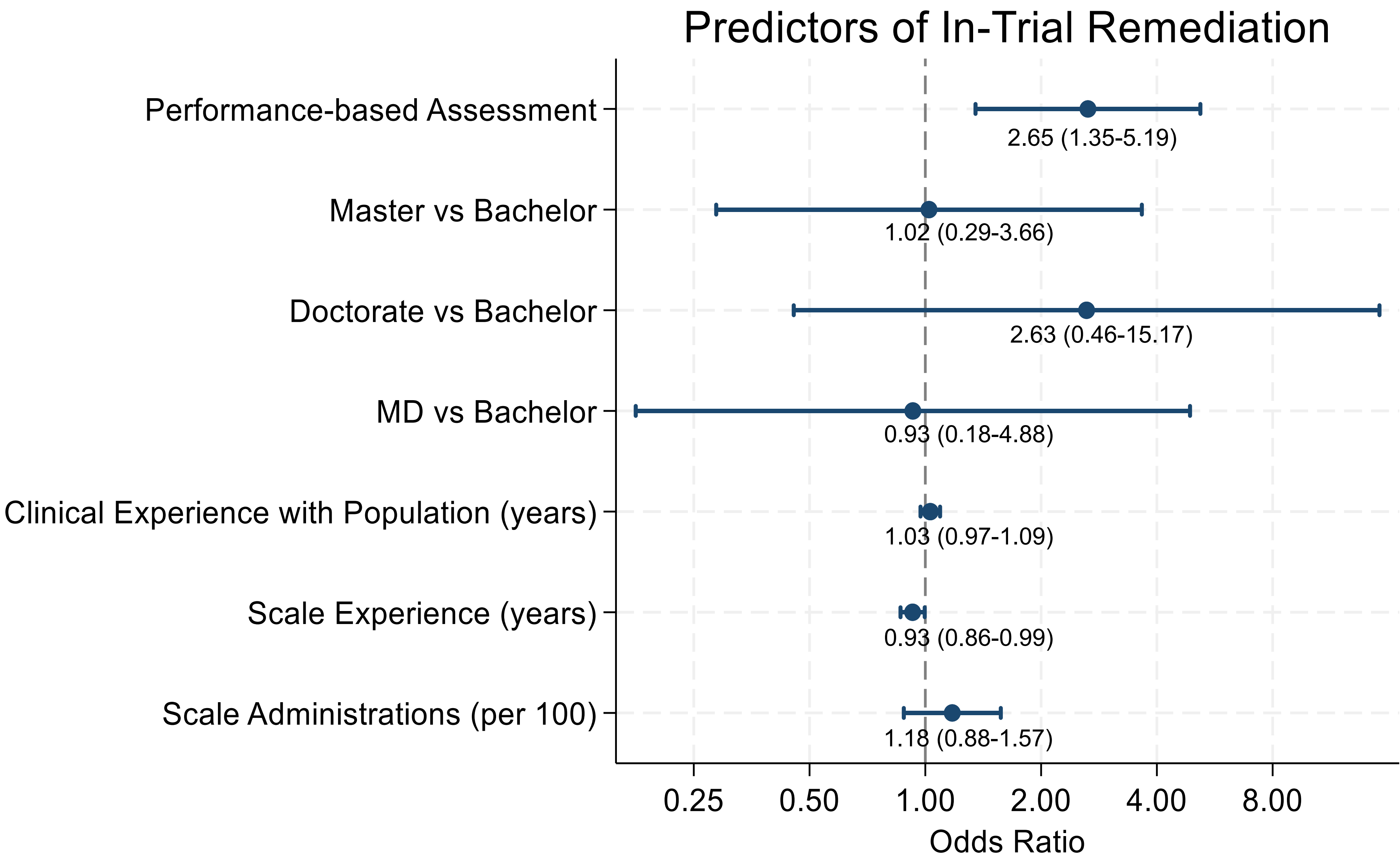
Results

Figure 1: Descriptives

- Performance-based assessment**, coded as 1 or more than 1 attempt (more than 1 attempt indicates that initial attempt did not meet pre-established quality threshold)
- Education**, dummy coded into 4 levels (with roughly comparable levels collapsed within a level)
- Clinical Experience with Population**, reported as # years of clinical experience with individuals with Alzheimer’s Disease
- Scale Experience** (years), reported as # years of experience with respective scale
- Scale Administrations** (per 100), reported as number of recent administrations with respective scale (in last 12 months); for purposes of this analysis, the variable was scaled such that a one-unit increase aligns with 100 additional assessments reported in the past year

Variable	N	M	SD
Education (dummy coded)			
Bachelor	70		
Master	184		
Doctorate	63		
Medical Doctor	160		
Clinical Experience with Population		12.54	8.81
Scale Experience		8.80	6.92
Scale Administrations		82.12	143.17

Figure 2: Mixed-effects Logistic Regression conducted in Stata 19. The full model also accounted for variation across trials and scale utilized (not reported here). The overall model result was statistically significant, $\chi^2(13) = 25.54$, $p = .02$. All Confidence intervals are 95%.



References

Patrick, K., Farrell, M., Sirbu, C., Karazsia, B., McNamara, C. (2025, July) [Testing Industry Assumptions About Raters’ Clinical Experience and Performance During Alzheimer’s Disease Clinical Trials](#). Poster presented at AAIC, Denver, CO.

Kobak, K. A., Engelhardt, N., Williams, J. B., & Lipsitz, J. D. (2004). Rater training in multicenter clinical trials: issues and recommendations. *Journal of Clinical Psychopharmacology*, 24, 113-117.



IQVIA Rater Services, formerly Cronos Clinical Consulting Services, Inc.