

ESOURCE ADMINISTRATION OF THE CDR: PRELIMINARY VALIDATION OF INTERNAL CONSISTENCY CHECKS

Christopher Randolph^{1,2}, Christopher Weber¹, Lori Garzio¹, Selam Negash¹, Peter Böhm¹
¹ MedAvante, Inc. ² Loyola University Medical Center

INTRODUCTION

- The Clinical Dementia Rating scale (CDR) is widely used as sole primary and co-primary endpoint in therapeutic clinical trials of Alzheimer’s disease (AD).
- However, the CDR is challenging to score and scoring errors are common (Tractenberg, Schafer, Morris, 2001; Rockwood et al., 2000).
- We recently developed a tablet-based electronic source (eSource) data capture and monitoring investigative study platform with built-in consistency checks (“flags”) to improve scoring reliability.
 - The consistency checks are based on extensive training experience gained through several thousand expert reviews of CDR assessments.
 - The eSource platform can trigger interventions on many items in the CDR by providing raters with real-time queries and cross-checks prior to finalizing scores.

The goal of this study was to validate such internal consistency checks by examining:

- 1) How often flags would have been triggered in paper-based administration of the CDR;
- 2) How often the alerts were associated with scoring errors.

METHODS

- The CDR is a semi-structured interview of the subject and an informant to characterize cognitive and functional changes associated with AD and dementia (Morris, 2003).
 - The scale assesses six domains: Memory, Orientation, Judgment & Problem Solving, Community Affairs, Home & Hobbies, and Personal Care.
- In the present study, a sample of paper-based CDR assessments was randomly selected from a recent clinical trial of mild-to-moderate AD.
 - The sample consisted of 200 CDR assessments completed by a total of 110 raters at 94 sites in 11 countries.
- Consistency checks were retrospectively applied to each of the paper-based assessments to determine how often flags would have been triggered if they had been available during scoring to alert raters.
 - For example, a box-score of 0 or 0.5 in the Memory domain would trigger a flag if an informant responded “rarely” to the question, “Can he recall recent events?”
- CDR assessments that would have triggered any flags were then cross-checked against scoring by a trained and calibrated central cohort of reviewers to identify any scoring discrepancies.

DISCUSSION

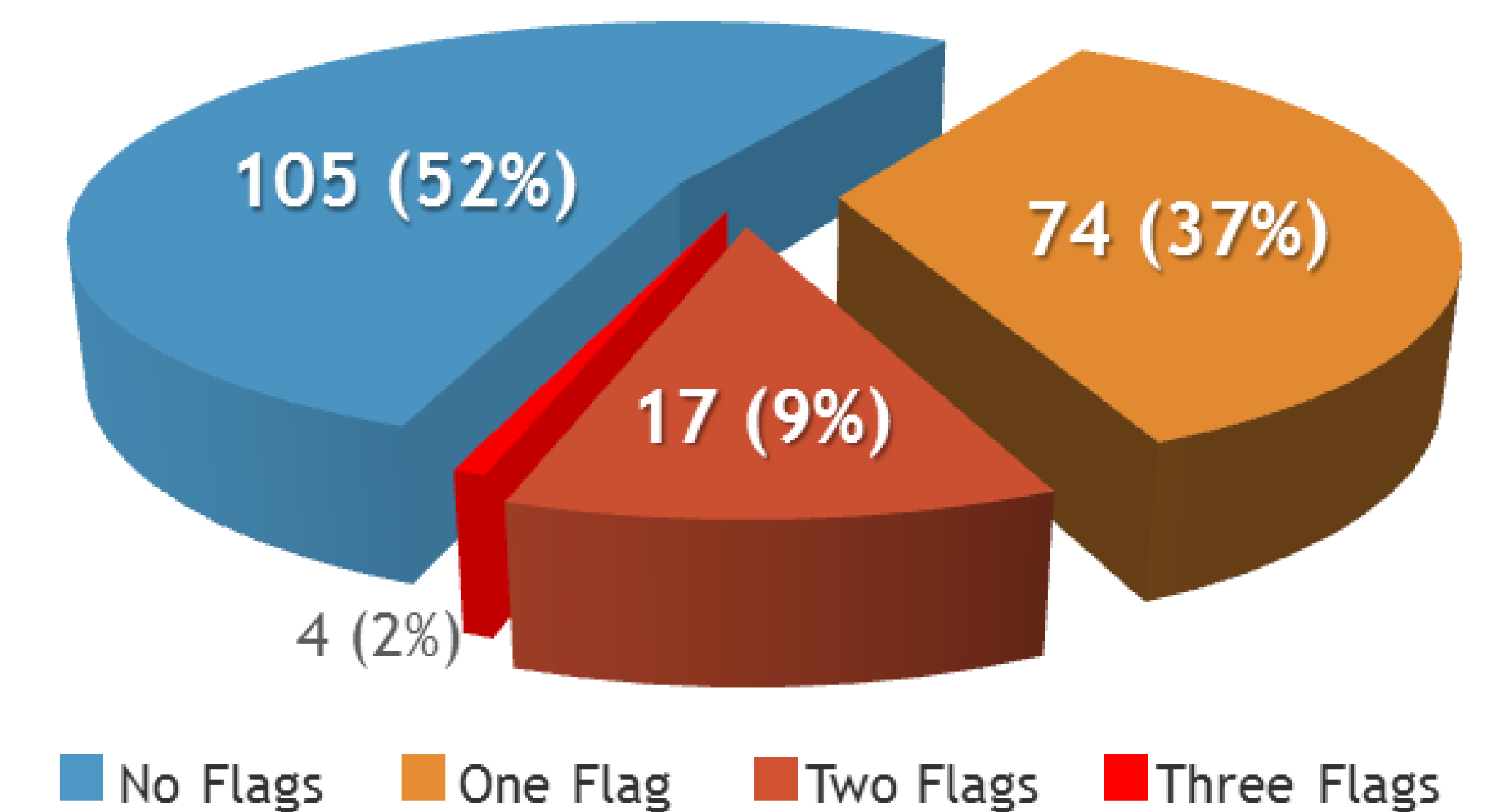
- The consistency checks built into the eSource platform would have been triggered for raters on nearly 50 percent of paper-based CDR administrations.
- The flags would have alerted raters to scoring discrepancies in more than 60 percent of the CDR administrations.
- The consistency checks are effective at identifying scoring discrepancy in domains that are particularly difficult to score, such the Memory domain (Tractenberg, et al., 2001).

REFERENCES

Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*. 1993 Nov; 43(11):2412-4. PubMed PMID: 8232972
 Rockwood K, Strang D, MacKnight C, Downer R, Morris JC. Interrater reliability of the clinical dementia rating in a multicenter trial. *American Geriatrics Society*. 2000 48:558-559. PMID: 10811551
 Tractenberg RE, Schafer K, Morris JC. (2001). Interobserver disagreements on clinical dementia rating assessment: interpretation and implications for training. *Alz Dis Assoc Disord*. 2001 Jul-Sep; 15(3):155-61. PMID: 11522933

RESULTS

FIGURE 1. CDR assessments with flags



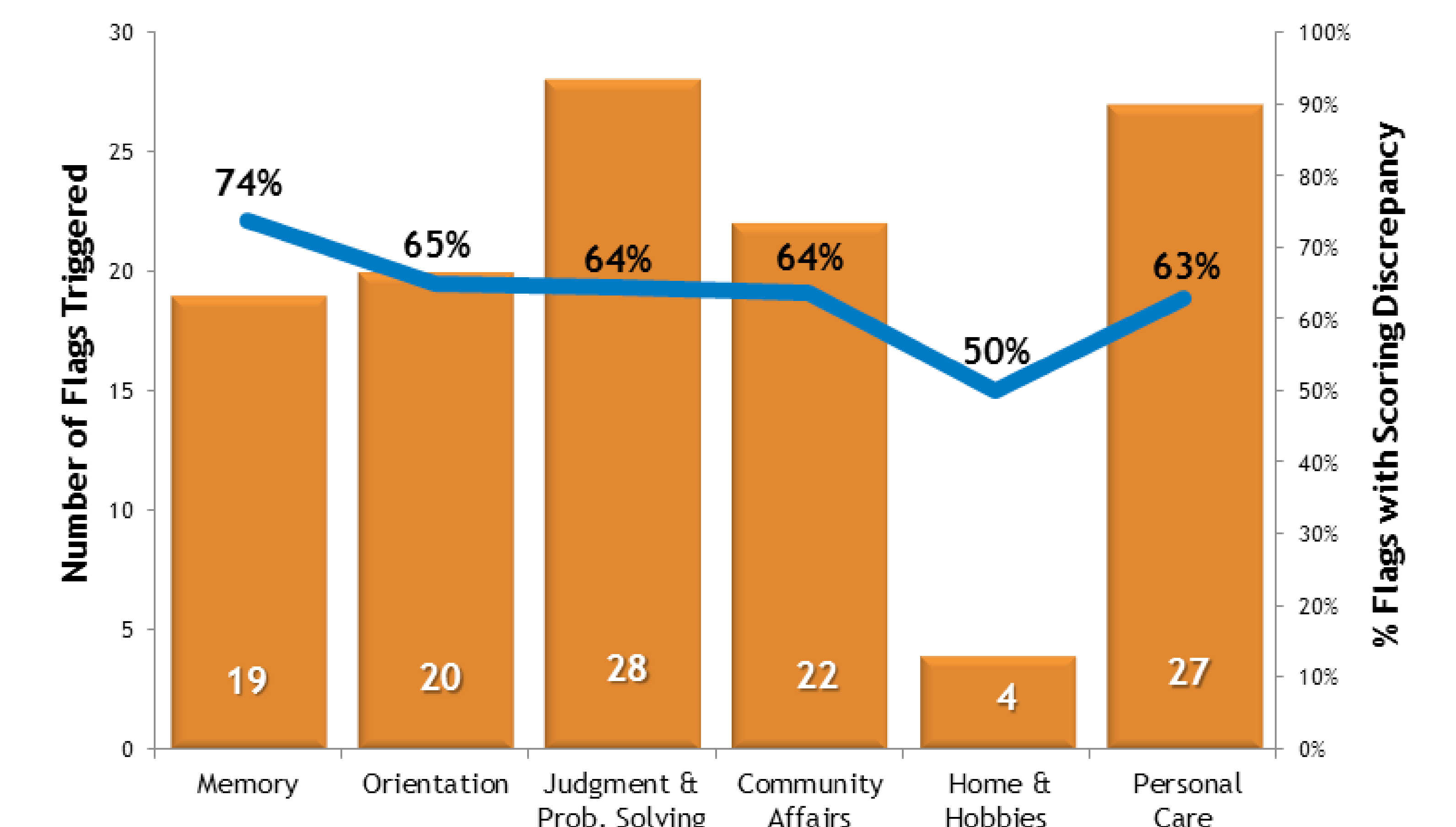
- 95 (47.5 percent) of the CDR assessments would have triggered at least one flag. (Figure 1)
 - 11 percent would have triggered two or more flags.
- Of the assessments with flags, 63 percent contained scoring discrepancy.

- The number of flags triggered at domain level (orange bar) along with the percentages associated with a scoring error (blue line) are displayed in Figure 2.

- The number of flags triggered within a domain ranged from 28 (Judgment & Problem Solving) to four (Home & Hobbies).

- For all domains, 50 percent (or more) of the flags triggered were associated with a scoring error.
 - For example, in the Memory domain, 14 out of the 19 flags (74 percent) contained scoring discrepancy.

FIGURE 2. Flags and scoring discrepancies within CDR domains



CONCLUSION

- An eSource platform with multi-level clinical guidance for CDR administration can reduce scoring errors that contribute to poor interrater reliability, thereby improving signal detection.

