RISK-BASED ERROR PREVENTION USING VIRGIL FLAGS IN PANSS ADMINISTRATION AND SCORING

Negash S[1], Stein L[1], Prochnik E[1], Webber-Lind B[1], Detke MJ[1,2], Williams JBW[1,3]

[1] MedAvante, Inc. [2]Indiana University School of Medicine [3]Depts. of Psychiatry and Neurology, Columbia University College of Physicians and Surgeons

THE METHODOLOGICAL QUESTION BEING ADDRESSED

Risk-based monitoring (RBM) has recently been recommended by the FDA[1] as an alternative approach to 100% source data verification of all trials and is currently being pursued in various clinical trials to help improve sponsor oversight efficiency and ensure study quality. The extent to which this type of targeted, risk-based approach identifies scoring errors in measures of efficacy in clinical trials, however, is less understood.

INTRODUCTION

- Sponsors of clinical trials seeking ways to improve oversight efficiency and ensure the quality of study data are adopting technology-based solutions such as centralized risk-based monitoring. Another aspect of ensuring data quality is improving accuracy in the scoring of rating scales, a goal that also lends itself to a risk-based technology solution capable of identifying scoring errors in efficacy measures before the data are submitted.
- The Positive and Negative Syndrome Scale (PANSS)[2] is a complex scale prone to administration and scoring errors that can contribute to poor interrater reliability and inaccurate clinical trial results.
- The Virgil Investigative Study Platform collects electronic source (eSource) data, prevents calculation errors and highlights missing data for raters as they administer and score assessments to standardize measurements and improve data quality. For PANSS items, this electronic clinical outcome assessments (eCOA) platform also triggers interventions to provide raters with clinical guidance in real time, including "flags" to alert the rater to scoring inconsistencies.
- In the present study, we examined scoring errors in paper-based administration of the PANSS in schizophrenia trials to determine the extent to which occurrences of the flags are, in fact, associated with error rates in scoring of the PANSS.

METHODS

- The Virgil eCOA platform has 18 flags built in to identify potential risk of scoring errors in the PANSS. These include pairs of incompatible item scores within a visit, such as:
  - A rating on P1 indicating absent or minimal delusions, along with a rating on P5 of clear cut delusions
  - A rating on G2 indicating absence or denial of anxiety, along with verbal report of anxiety on G4
- We also examined four types of between-visit inconsistencies, such as same response on all items from a previous visit.
- These flags were then applied to paper-based PANSS assessments aggregated from eight randomized, double blind, placebo-controlled schizophrenia trials.
- 768 paper-based assessments were reviewed by the same central cohort of blinded clinicians who identified scoring errors by reviewing recordings and worksheets from each assessment.
- The scoring errors were compared against the Virgil flags to determine how often a flag would have been triggered in assessments that had two or more errors.

RESULTS

- Figure 1 shows the percentages of PANSS assessments that would have triggered flags.
  - 422 (55%) assessments would have triggered two or more flags. Most assessments would have triggered one or two flags (68%), with assessments triggering more than three flags likely to be low (5%).
  - Because some assessments had more than one alert, a total of 1,297 individual alerts would have been triggered in this cohort.
- Figure 2 indicates the scoring errors associated with the flags.
  - 565 assessments (74%) had two or more scoring errors, suggesting that scrutiny of PANSS ratings scale administration is essential.
  - Of those that triggered one or more flags (n = 678), 73% were identified in central review as having scoring errors.
  - In assessments that had scoring errors, the flags that occurred with heightened frequency included:
    - Depressed or low mood rated 2 or higher without verbal confirmation from the subject (71%)
    - Inconsistencies between distrustfulness and active social avoidance (33%)
    - Inconsistencies between lack of insight and delusions (17%)
    - Between-visit total score change of greater than 25% (12%)

FIGURE 1.
Percent Assessments Triggering PANSS Flags

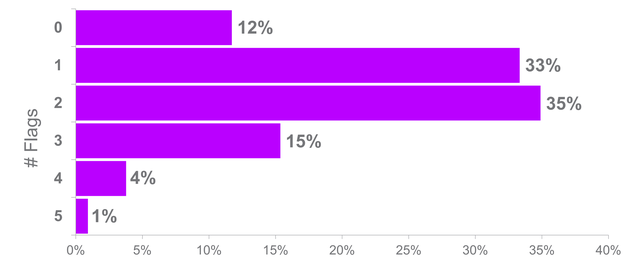
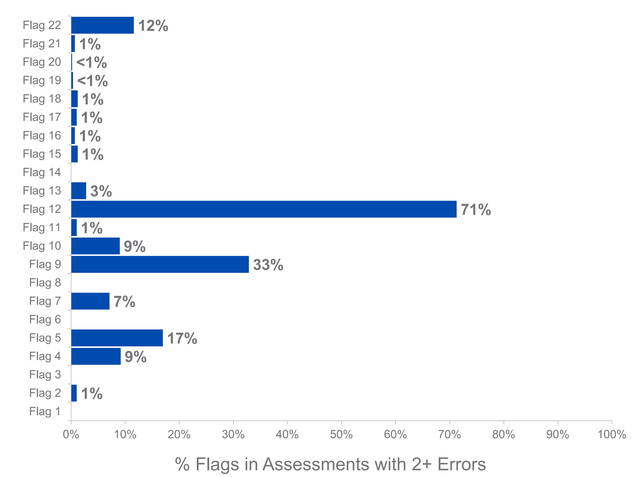


FIGURE 2.
Percent Flags in Assessments with Two or More Errors



CONCLUSIONS

- Risk-based analysis is useful in identifying potential scoring errors, and real-time technology can assist in preventing problematic ratings by directing the rater's focus to the most critical data elements that need attention.
- It is important to select an appropriate number of alerts – too many may diminish consideration by raters; too few may miss inaccuracies that affect subscale and total scores.
- Not all scoring alerts indicate a scoring error, but should prompt raters to reconsider conflicting ratings in light of scoring conventions and potential clinical inconsistencies.
- The Virgil eCOA platform with real-time clinical guidance, scoring anchors, and item descriptors minimizes scoring inconsistencies to reduce error variance and improve signal detection.

References
[1] U.S. FDA, Oversight of Clinical Investigations – A Risk-Based Approach to Monitoring Guidance, Draft (August 2011)
[2] Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophrenia Bulletin 1987 13 (2): 261–76.

The authors report no conflicts of interest for this work.

©2016 MedAvante, Inc.

MedAvante®