Early Indicators of Poor Data Quality in Schizophrenia Clinical Trials

Kott, Alan¹; Daniel, David G.²
¹Bracket, Prague, Czech Republic; ²Bracket, McLean, VA, USA

METHODOLOGICAL QUESTION
We intend to examine whether the presence of data quality issues after randomization is associated with the presence of data quality issues identified during the screening period

INTRODUCTION
A primary focus of risk-based data quality monitoring program in clinical trials is early identification and remediation of issues that may detract from signal detection

■ Ideally such a program should identify problematic ratings before a subject is randomized into the study

■ In the current analysis, we examined whether the incidence of data quality issues after randomization was associated with the presence of data quality issues in the screening period

METHODS
■ We analysed blinded data from 14 international double blind placebo controlled schizophrenia trials involving 10,056 subjects (67,584 visits) using univariate negative binomial regression

■ The incidence of data quality issues after randomization was compared between those subjects who recorded a data quality issue in the screening phase and those who did not

■ The examined data quality indicators are listed in table 1

Table 1: List of data quality indicators
<table>
<thead>
<tr>
<th>Data Quality Indicator</th>
<th>Possible hit (randomized)</th>
<th>Subtext graphics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any data quality concern</td>
<td>Yes</td>
<td>Any Data Quality Concern</td>
</tr>
<tr>
<td>Identical scoring of ECO/CGI-PANS items across consecutive visits</td>
<td>Yes</td>
<td>Identical PANSS</td>
</tr>
<tr>
<td>Identical PANSS Logical Inconsistencies</td>
<td>Yes</td>
<td>Large Change</td>
</tr>
<tr>
<td>On the Incidence of Post-randomization Data Quality Concerns</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

■ We fitted all possible univariate models (47) testing for the effect of presence of data quality concerns before randomization on the incidence of data quality concerns after randomization

RESULTS

■ We identified significant effect of the presence of data quality concerns in the screening phase on the incidence of the same data quality concerns after randomization (figures 1 - 4)

■ Other significant findings include:

■ The presence of any data quality concern before randomization significantly increased the incidence of all examined data quality concerns after randomization except of discrepancies between PANSS change from baseline and CGI-I scores (figure 1)

■ The presence of 30/30 PANSS items identical at screening and baseline significantly decreased the incidence of post-randomization erratic ratings (figure 2)

■ The presence of large changes in the PANSS from screening to baseline significantly increased the incidence of post-randomization erratic ratings (figure 3)

■ The presence of logical inconsistencies among individual PANSS items before randomization significantly increased the incidence of large and erratic changes in the PANSS and decreased the post-randomization incidence of PANSS identical ratings (figure 4)

■ We identified significant effect of the presence of pre-randomization data quality concerns on the incidence of the same and related data quality concerns after randomization

■ This represents an important finding because it indicates that raters who commit serious rating and interviewing errors in the screening phase are likely to commit significantly more errors later in the study with the potentially negative effect on signal detection

■ The findings underline the necessity to intervene early to prevent the identified raters to repeat the errors

■ Many of the examined data quality concerns could be prevented with the potential of immediate error detection and feedback features, and in-study audio/video monitoring (Kott & Daniel, 2015)

REFERENCES