Methods for Responder Analysis of Patient-Reported Outcomes

Alexandra I. Barsdorf, Ph.D.1, Kelly H. Zou, Ph.D.1, Martin O. Carlsson, M.S.1 and Joseph C. Cappelleri, Ph.D.2

Pfizer Inc., 1New York, N.Y. 10017 and 2Groton, C.T. 06340, U.S.A.

1. Introduction

- Responder analysis is about within-group change over time, and clinically important difference (CID) is about between-group change.1,2
- Both are useful for analyzing a patient-reported outcome (PRO) meaningful to a patient.
- The FDA recommends plotting cumulative distribution functions (c.d.f.) to help interpret the clinical relevance of a PRO.3

2. Objectives

- To compare distribution- and anchor-based methods for the CID and evaluate the influence of correlation between the target PRO and the anchor on the resulting CID estimates.
- To combine responder analysis (within-individual change) and clinical significance (between-group difference) in a hybrid approach that involves between-group differences and within-individual changes to examine the CID for the PRO of interest.

3. An Example from Published Literature

- The outcome variable was the absolute change in total patient-reported pain intensity between two different time points.4 The observed scores ranged between -6 and 10. See Table 1.
- The quasi-continuous anchor variable GRS9 included 9 categories (1-9). They were dichotomized into 3 ordinal categories (‘importantly improved’, ‘no change’ vs. ‘negative change’) or dichotomized into 2 binary response types by combining ‘no change’ and ‘negative change’ as ‘otherwise’.5 Table 1. Descriptive Statistics of the Absolute Change of Marker, by Dichotomized Anchor GRS9

<table>
<thead>
<tr>
<th>Trichotomized Anchor Xa</th>
<th>Patient’s Condition</th>
<th>Sample Size</th>
<th>Mean (SD)</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 if GRS&gt;5</td>
<td>Importantly Improved</td>
<td>346</td>
<td>4.7(2.11)</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>1 to 10</td>
</tr>
<tr>
<td>0 if GRS=5</td>
<td>No Change</td>
<td>139</td>
<td>5.3(1.96)</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-2 to 4</td>
</tr>
<tr>
<td>-1 if GRS&lt;5</td>
<td>Negative Change</td>
<td>50</td>
<td>-0.1(-1.44)</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>-4 to 2</td>
</tr>
</tbody>
</table>

4. Methods

4.1 Distribution-Based Method

- A distribution-based method derives the CID as a score corresponding to Cohen’s benchmarks effect sizes (ES) 0.2, 0.5 or 0.8, with CID= ES SD(Y).
- SD(Y) is the pooled standard deviation at baseline regardless of the anchor category in X. See Table 2.

4.2 Anchor-Based Method

4.2.1 Anchor-Based Method: 1 Sample Means

- Conventionally, the anchor-based CID is the difference between sample means of Y by response category, CID=Mean(R)-Mean(NR)=∆R-∆NR.

4.2.2 Anchor-Based Method: 2 ROC Analysis

- Alternatively, a distribution-free anchor-based method is based on a receiver operating characteristic curve (ROC) analysis to differentiate between the two response categories.6
- Using Youden’s index (YI), where YI=Specificity-Sensitivity×1 =F1(∆X|X=R)×[1-G1(∆X|X=NR)]+F1(∆X|X=NR)×[1-G1(∆X|X=R)], and F and G are the c.d.f.s for NR and R categories by dichotomized anchor. Fig. 1. An illustration of the Anchor-Based Method Using Maximal Youden’s Index for CID Estimation

Table 3. An Illustration of the Anchor-Based Method Using Sample Means for CID Estimation

<table>
<thead>
<tr>
<th>R Category: &gt;0 Point Change</th>
<th>Minimum Responder Subset vs. No Change in Outcome Measure</th>
<th>Estimated CID on Outcome Measurement Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Category: &gt;1 Point Change</td>
<td>Mean Change in Outcome Measure</td>
<td>Estimated CID</td>
</tr>
<tr>
<td>NR Category: 0 Point Change</td>
<td>Mean Change in Outcome Measure</td>
<td>Estimated CID</td>
</tr>
</tbody>
</table>

CID Method Effect Size ES 0.1 0.2 0.5 0.7 1.0

Small (0.2) 0.80 (1.65,3.08) 0.80 (1.65,3.08) 0.80 (1.65,3.08) 0.80 (1.65,3.08) 0.80 (1.65,3.08)
Medium (0.5) 0.50 (1.35,2.52) 0.50 (1.35,2.52) 0.50 (1.35,2.52) 0.50 (1.35,2.52) 0.50 (1.35,2.52)
Large (1.0) 0.25 (0.87,2.57) 0.25 (0.87,2.57) 0.25 (0.87,2.57) 0.25 (0.87,2.57) 0.25 (0.87,2.57)

5. Results

- Table 4 compares the estimated CIDs, with 95% confidence intervals (CI), for change in pain intensity.7 These CIDs are variable across methods.

6. Monte-Carlo Simulations

- Generate 1000 pairs of (Y1, Y2)=BN(µ1, 0.5), (10, 10), which are moderately correlated (0.5) between t0 and tT. Let the baseline CIDs be Y1=Y2.
- Let (X,∆)=BN(1.0, 1.0), with the anchor and change score having means 1 and 10, and various correlations ρ between these two variables.
- Based on 1000 replicates, report the median CIDs (95% CIs). See Table 5. Table 5. Estimated CIDs based on Monte-Carlo sampling of simulated longitudinal data

CID Method Effect Size ES 0.1 0.2 0.5 0.7 1.0

Small (0.2) 0.80 (1.65,3.08) 0.80 (1.65,3.08) 0.80 (1.65,3.08) 0.80 (1.65,3.08) 0.80 (1.65,3.08)
Medium (0.5) 0.50 (1.35,2.52) 0.50 (1.35,2.52) 0.50 (1.35,2.52) 0.50 (1.35,2.52) 0.50 (1.35,2.52)
Large (1.0) 0.25 (0.87,2.57) 0.25 (0.87,2.57) 0.25 (0.87,2.57) 0.25 (0.87,2.57) 0.25 (0.87,2.57)

7. Conclusions

- CIDs vary across methods and correlations between outcome and anchor.
- A method for CID should be accompanied by a c.d.f. plot to assess clinical meaningfulness of a PRO across the entire range of responses.1,2

8. References

Figure 1. An Illustration of the Anchor-Based Method Using Maximal Youden's Index for CID Estimation

- NR = Non-Responders
- R = Responders
- CID = Max(Youden Index)

- Sensitivity = $P(Y > \text{Threshold} | X = R)$
- Specificity = $P(Y \leq \text{Threshold} | X = NR)$
Figure 2. ROC Curve (Improved vs. Otherwise)
Figure 3.A C.D.F. Plot for Improved vs. Otherwise (No Change or Negative Change)
Figure 3.B C.D.F. Plot for Improved, No Change vs. Negative Change