Sensitivity of PROC DISCRIM for Different List of Variables to Separate a Study Population by Treatment Subgroups in Clinical Trial With a New Antidepressant

Lev Sverdlov, Ph.D., Innapharma, Inc., Park Ridge, NJ

ABSTRACT
Data were evaluated from a double blind, placebo-controlled, pilot clinical trial (N=52) with a new antidepressant. Discriminant analysis (PROC DISCRIM) was used to separate the drug-treated from placebo populations by treatment subgroups. As suggested by clinical psychiatrists, two different lists of variables were tested to check the sensitivity of discriminant analysis to the clinical assessments. The first list of variables in PROC DISCRIM included 7 primary and secondary clinical assessments for two time points: end of treatment and peak effect. The second list included just one primary clinical assessment for 7 time points from the first day of treatment to the end of observation. Separation with PROC DISCRIM was effective for both lists of variables with the overall correct classification by treatment subgroups from 80% to 83%. All of the methodological aspects and the SAS® code can be used in some pivotal trials in the CNS therapeutic area.

1. INTRODUCTION
We have previously reported [1] the use of discriminant analysis (PROC DISCRIM) to separate a drug-treated from a placebo population after treatment with a new antidepressant [2]. This approach was developed to provide additional methodology for dealing with common statistical analysis of clinical data [3,4] and was effective in separating the drug-treated population from placebo in the pilot clinical trial. We now report the use of the same approach with two different lists of variables for another trial to test the sensitivity of this statistical procedure to the clinical assessments.

2. METHOD
2.1. Study Design
The study design has been previously described in the paper [5]. Fifty-two male and female physically healthy subjects, 18 years or older, diagnosed with non-psychotic major depression were enrolled in this pilot phase 2 study (inpatient, double-blind, randomized, placebo-controlled, parallel-design, single-center). The study investigated the efficacy and safety of a new drug administered subcutaneously for the 5-day treatment cycle. At the end of the 5-day treatment cycle, the subjects returned to the facility once weekly for the next 4 weeks for follow-up evaluations. The placebo group included 26 subjects and the drug group also included 26 subjects. The primary efficacy assessment was the 21-item Hamilton-Depression Rating Scale (HAMD) (change and % change from baseline). The secondary efficacy assessments were Montgomery-Asberg Depression Rating Scale (MADRS), Carroll Self-Rating Scale (CSRS) and Clinical Global Improvement (CGI). The last-observation-carried-forward method (LOCF) was used for missing data. A retrospective pharmacokinetic analysis permitted the definition of the Minimum Projected Therapeutic Concentration (MPTC) and the subsequent division of the study population into three treatment subgroups: placebo subgroup (all subjects from the placebo group), drug-treated subgroup 1 with plasma drug concentrations in the therapeutic range (above MPTC) and drug-treated subgroup 2 with plasma drug concentrations below the therapeutic range (below MPTC).

2.2. Statistical Evaluation (Discriminant Analysis)
SAS/STAT® [6,7] is a powerful tool for discriminant analysis with some options allowing selection of parametric or non-parametric methods, linear or quadratic classified functions, equal or unequal prior probability for each level of classification variable, with or without calculation of new variables with canonical scores, et al. The first list of variables for discriminant analysis included psychometric assessments (primary and secondary) for two time points: day 7 (the first evaluation after treatment) and day 14 (peak effect of response). It was the same list of variables we previously used for cluster analysis [8]. The second list of variables was suggested by clinical psychiatrists and included only the primary clinical assessment for 7 time points from the first day of evaluation to the end of observation. The total number of variables for both options was equal to 7. Because of the relatively small sample size in this study, there was no replication group and all of the subjects were included only in the training group.
2.3. SAS Code

Macro 1: Discriminant Analysis for Training Data Set

```sas
%macro da_train (dataset_1, var_1, var_2, var_3, var_4, var_5, var_6, var_7);
proc discrim data=&dataset_1 anova distance listerr
   method=normal
   out=testout
   outstat=teststat
   outd=outd;
   class plevel;
priors proportional;
var &var_1 &var_2 &var_3 &var_4
   &var_5 &var_6 &var_7;
title 'Discriminant analysis for 7 clinical assessments for Training data set';
run;
proc print data=testout;
var subjid subjinit treat plevel
   &var_1 &var_2 &var_3 &var_4
   &var_5 &var_6 &var_7
   P H L _into_;
run;
proc print data=outd;
var subjid subjinit treat plevel
   &var_1 &var_2 &var_3 &var_4
   &var_5 &var_6 &var_7
   P H L ;
run;
%mend da_train;
/* dataset_1 - training data set             */
/* var_1, var_2 var_3, var_4, var_5, var_6, var_7 - seven clinical assessments             */
/* plevel - classification variable         */
```

Macro 2: Replication Data Set

```sas
%macro da_replc (dataset_1, dataset_2, var_1, var_2, var_3, var_4, var_5, var_6, var_7);
proc discrim data=&dataset_1
   testdata=&dataset_2 anova distance listerr
   method=normal
   out=testout
   outstat=teststat
   outd=outd;
   class plevel;
priors proportional;
var &var_1 &var_2 &var_3 &var_4
   &var_5 &var_6 &var_7;
title 'Discriminant analysis for 7 clinical assessments for Replication data set';
run;
%mend da_replc;
/* dataset_2 - training data set             */
```

3. RESULTS

Of the 52 enrolled and randomly assigned subjects (i.e., intent-to-treat data set), 49 subjects (94.2%) completed the initial treatment (i.e., evaluable data set). The MPTC for pharmacokinetic evaluation was defined as 5.0 ng/mL at 1 hr after dosing [5], and for subjects treated with drug this value was used to select subgroup 1 (10 subjects) and subgroup 2 (15 subjects) with plasma drug concentrations above or below the therapeutic range. A summary of discriminant analysis and discriminant classification accuracy for all separations is presented in sections 3.1 – 3.4.

Option 1: The first list of variables

3.1. Separation of Subgroup 1 from Placebo

Seven of 10 subjects (70.0%) from subgroup 1 and 17 of 24 subjects (70.8.7%) from the placebo group were classified correctly and confirmed the prior known actual subgroups. The error rate of classification was 29.4%, which means 70.6% of correctly classified subjects were from both subgroups combined (see Table 1).

3.2. Separation of Subgroup 1 from Subgroup 2

Nine of 10 subjects (90.0%) from subgroup 1 and 13 of 15 subjects (86.7%) from subgroup 2 were classified correctly and confirmed the prior known actual subgroups. The error rate of classification was 12.0%, which means 88.0% of correctly classified subjects were from both subgroups combined (see Table 2).

Option 2: The second list of variables

3.3. Separation of Subgroup 1 from Placebo

Eight of 10 subjects (80.0%) from subgroup 1 and 20 of 24 subjects (83.3%) from the placebo group were classified correctly and confirmed the prior known actual subgroups. The error rate of classification was 17.6%, which means 82.4% of
correctly classified subjects were from both subgroups combined (see Table 3).

3.4. Separation of Subgroup 1 from Subgroup 2

Nine of 10 subjects (90.0%) from subgroup 1 and 12 of 15 subjects (80.0%) from subgroup 2 were classified correctly and confirmed the prior known actual subgroups. The error rate of classification was 12.0%, which means 88.0% of correctly classified subjects were from both subgroups combined (see Table 4).

4. DISCUSSION

Figure 1 is a comparable summary for the effect of separation with PROC DISCRIM for both lists of variables. There was no significant difference in the results of separation between Option 1 and Option 2.

In contrast, the traditional currently used analysis of longitudinal data did not confirm the effect of separation between treatment subgroups on the basis of the single endpoint at the end of observation [5]. It is important to mention that after discriminant analysis the majority of responders were classified to subgroup 1 (above MPTC) versus the majority of nonresponders to subgroup 2 (below MPTC).

While our findings summarize the results from a pilot study with a limited sample size, discriminant analysis taking into account concentration of drug in plasma from pharmacokinetic evaluation successfully created a very comprehensive picture of the separation of the drug-treated population from the placebo population. We are now planning to use discriminant analysis to evaluate the results of future pivotal clinical studies.

5. CONCLUSIONS

Discriminant analysis (PROC DISCRIM) was very effective in separating a heterogeneous study population of subjects diagnosed with major depression into three treatment subgroups using HAMD-21 scores for all the available time points. Discriminant analysis demonstrated that the overall correct classification by treatment subgroups for both lists of variables was from 82.4% to 86.7%. The results indicate that multivariate discriminant analysis is more reflective of the dynamics of drug effect than assessment at a single endpoint (particularly for atypical dose regimens) and provides a valid additional statistical approach to support conclusions of efficacy. All of the methodological aspects and SAS codes presented in this paper can be used during drug development in some pivotal studies in the CNS therapeutic area. These techniques allow for a greater sample size, while separating the study population into training (calibration) and replication groups.

REFERENCES


TRADEMARKS

SAS® is a registered trademark of SAS Institute, Inc. in the USA and other countries. ® indicates USA registration.

CONTACT INFORMATION

Lev Sverdlov, Ph.D., Innapharma, Inc.
Phone: 201-505-1549, Ext. 645
E-mail: LSVERDLOV@AOL.COM
Table 1. Discriminant Analysis (the first list of variables) for the Separation Population with Plasma Drug Concentrations Above the Therapeutic Range (Subgroup 1) Versus Placebo

<table>
<thead>
<tr>
<th>Actual Treatment Subgroup</th>
<th>Number of Subjects</th>
<th>Classification After Discriminant Analysis, Number (%) of Subjects by Treatment Subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subgroup 1</td>
</tr>
<tr>
<td>Subgroup 1 (Above MPTC)</td>
<td>10</td>
<td>7 (70.0%)</td>
</tr>
<tr>
<td>Placebo</td>
<td>24</td>
<td>7 (29.2%)</td>
</tr>
<tr>
<td>Total</td>
<td>34</td>
<td>14 (41.2%)</td>
</tr>
<tr>
<td>Total Number of Subjects with Correct Classification:</td>
<td>(24 =7+17) of 34 (70.6%)</td>
<td></td>
</tr>
<tr>
<td>Error Rate of Classification:</td>
<td>(10 =3+7) of 34 (29.4%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Discriminant Analysis (the first list of variables) for the Separation Population with Plasma Drug Concentrations Above the Therapeutic Range (Subgroup 1) Versus Below the Therapeutic Range (Subgroup 2)

<table>
<thead>
<tr>
<th>Actual Treatment Subgroup</th>
<th>Number of Subjects</th>
<th>Classification After Discriminant Analysis, Number (%) of Subjects by Treatment Subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subgroup 1</td>
</tr>
<tr>
<td>Subgroup 1 (Above MPTC)</td>
<td>10</td>
<td>9 (90.0%)</td>
</tr>
<tr>
<td>Subgroup 2 (Below MPTC)</td>
<td>15</td>
<td>2 (13.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>11 (44.0%)</td>
</tr>
<tr>
<td>Total Number of Subjects with Correct Classification:</td>
<td>(22 =9+13) of 25 (88.0%)</td>
<td></td>
</tr>
<tr>
<td>Error Rate of Classification:</td>
<td>(3 =1+2) of 25 (12.0%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Discriminant Analysis (the second list of variables) for the Separation Population with Plasma Drug Concentrations Above the Therapeutic Range (Subgroup 1) Versus Placebo

<table>
<thead>
<tr>
<th>Actual Treatment Subgroup</th>
<th>Number of Subjects</th>
<th>Classification After Discriminant Analysis, Number (%) of Subjects by Treatment Subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subgroup 1</td>
</tr>
<tr>
<td>Subgroup 1 (Above MPTC)</td>
<td>10</td>
<td>8 (80.0%)</td>
</tr>
<tr>
<td>Placebo</td>
<td>24</td>
<td>4 (16.7%)</td>
</tr>
<tr>
<td>Total</td>
<td>34</td>
<td>12 (35.3%)</td>
</tr>
<tr>
<td>Total Number of Subjects with Correct Classification:</td>
<td>(28 =8+20) of 34 (82.4%)</td>
<td></td>
</tr>
<tr>
<td>Error Rate of Classification:</td>
<td>(6 =2+4) of 34 (17.6%)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Discriminant Analysis (the second list of variables) for the Separation Population with Plasma Drug Concentrations Above the Therapeutic Range (Subgroup 1) Versus Below the Therapeutic Range (Subgroup 2)

<table>
<thead>
<tr>
<th>Actual Treatment Subgroup</th>
<th>Number of Subjects</th>
<th>Classification After Discriminant Analysis, Number (%) of Subjects by Treatment Subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subgroup 1</td>
</tr>
<tr>
<td>Subgroup 1 (Above MPTC)</td>
<td>10</td>
<td>9 (90.0%)</td>
</tr>
<tr>
<td>Subgroup 2 (Below MPTC)</td>
<td>15</td>
<td>3 (20.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>12 (48.0%)</td>
</tr>
<tr>
<td>Total Number of Subjects with Correct Classification:</td>
<td>(22 =9+13) of 25 (88.0%)</td>
<td></td>
</tr>
<tr>
<td>Error Rate of Classification:</td>
<td>(3 =1+2) of 25 (12.0%)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Effect of Separation (PROC DISCRIM) between Subgroup 1 (patients with plasma drug concentration at 1 h after dosing above MPTC) versus Placebo, and versus Subgroup 2 (patients with plasma drug concentration at 1 h after dosing below MPTC) as a function of variables for discriminant analysis (options 1 or 2)

- MPTC - The Minimum Projected Therapeutic Concentration of drug in Plasma (5.0 ng/mL at 1 h after dosing)
- Option 1 - Seven Primary and Secondary variables for 2 time points (end of treatment and a week later)
- Option 2 - One Primary variable for 7 time points (from the beginning of treatment to the end of observation)