Paper SP07-2009

Adhoc Analysis of Site By Treatment Interaction

Alan Davis, PharmaNet, Inc., Charlotte, NC

Abstract

In order to enroll patients more quickly, and to help establish efficacy in a new drug, the sponsor of the drug may have multiple investigators conducting a clinical trial. Each investigator enrolls and treats patients at a particular center or site. Such a clinical trial is considered a multicenter study. In the analysis of results from such a trial, it is expected that sites will contribute a small effect so that including site as a covariate in the analysis of treatment effects is a common means of increasing the power of a multi-center clinical trial. Controlling for Site and Site by Treatment effects in a linear model increases the power of analysis when those factors are found to contribute significantly to reducing the residual error of the model. There may be factors associated with site, such as patient demographics and medical history, as well as site characteristics which may not have been otherwise accounted for, which increase or decrease a response variable for all treatments or specific treatments. ICH Guideline E9[1] recommends that if treatment effects vary by site, this should be explained, requiring further analyses of an ad-hoc nature. These can follow two possible strategies: looking for sites which are outliers, and looking for sites which have commonalities with other sites. By starting with residuals of the site and treatment-only linear model, we assume that simple treatment effects are removed. The residuals are then averaged by site and treatment and displayed graphically to look for sites which are similar.

Keywords

Multi-Center, Multi-Site, Site-by-Treatment Interaction, Interaction, Linear Model, Residual Analysis

Introduction

A linear model of a response variable y may have three independent terms: treatment, site, and site-by-treatment. Site-by-treatment is an interaction between site and treatment and consists of an amount specific to each level of site and treatment[2]. In the linear model for a response variable, interaction is indicated when the interaction term, adjusting for other factors in the model, is large enough that its p-value is small. A criterion for this p-value may be 0.1 rather than the 0.05 alpha, since the value of detecting interaction is important, and the study may have low power in detecting interaction.

Interaction may have a biological basis, ie, certain sites inherently have a different treatment response from others, yielding potentially interesting conclusions and possibly further research to find the cause of response variability.

Alternatively, site and treatment interaction may be an artifact resulting from an error in treatment assignment, bias in evaluation, or other issue, irrelevant to the biological effect. For example, if some sites are not blinded to the treatment, there may be a bias in evaluating results in favor of the treated group. Another possibility is that certain sites may erroneously apply the wrong method of evaluation, which would skew their results one direction or another.

If site by treatment interaction is strong enough, it can be detected using a linear model, even if sample size per site is small. This will be shown in the example below.

Methods

Example 1: Pure Site By Treatment Interaction

In this example we generate a dataset for 20 sites and two treatment groups, treatment one (1) and treatment two (2), with up to five subjects per site. Treatments are unbalanced within site. Additionally, there is one site, site 1, which has results for treatment 2 only. Treatment is determined randomly (about half and half) and the response variable y is determined randomly. Treatment effect is 10.0 for treatment 1 at half the sites and 0.0 for treatment 1 at the remaining sites. Conversely, treatment effect for treatment 2 is 10.0 at those sites where treatment 1 effect is zero, and 0.0 at those sites where treatment 1 effect is 10. About 90 percent of the subjects are evaluable, so that there is variability in the number of subjects per site.

The code to generate the dataset is as follows:

```
data clin;
do site=1 to 20;
do nsubj= 1 to 5;
trt= int(1.5 + uniform(98987));  ** TRT variable is about 50/50;
if (trt =1 and site <11 ) or (trt=2 and site > 10 )then eff=0; else eff=10;
y= 75 + rannor(877877)*10 + eff;
** Output about 90% of these and no site 1 trt 1;
if uniform(544321)> .1  and not (site=1 and trt=1) then output;
end;
end;
run;
```

Using the following code, we analyze the variation of y into site, treatment, and site-by-treatment interaction:

```
proc glm data=clin; title Analysis Showing Significant Interaction;
class site trt;
model y=trt|site;
run;
```

The analysis of y shows significant site by treatment interaction and no treatment effect. The Type III sum of squares table is:

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	1	27.839491	27.839491	0.34	0.5622
site	19	1607.087774	84.583567	1.03	0.4422
site*trt	18	3248.612391	180.478466	2.21	0.0140

Given this result, we begin an ad-hoc analysis to explain the interaction. Since interaction is that which is not explained by simple main effects, by removing any site and treatment main effects we can plot the residuals and look for possible causes of the interaction.

To analyze residuals we execute the following:

```
Title  *** Residual Analysis ***;
proc glm data=clin; *First Remove Trt and Site Effect;
 class trt site;
 model y=trt site;
 output out=trtadj r=resid;
run;
* Calculate mean at each level of site and trt;
proc summary data=trtadj; class site trt;
 var resid;
 output out=resid(where=(_TYPE_=3)) mean=;
run;
******* Transpose to get data ready for plot;
proc transpose data=resid out=tdata;
 by site;
 id trt;
 var resid;
run;
proc plot data=tdata; plot _2*_1=site/href=0 vref=0;run;
```

The graph of residuals shows that the sites cluster in the high treatment-1 and low treatment 2 quadrant (Figure 1).

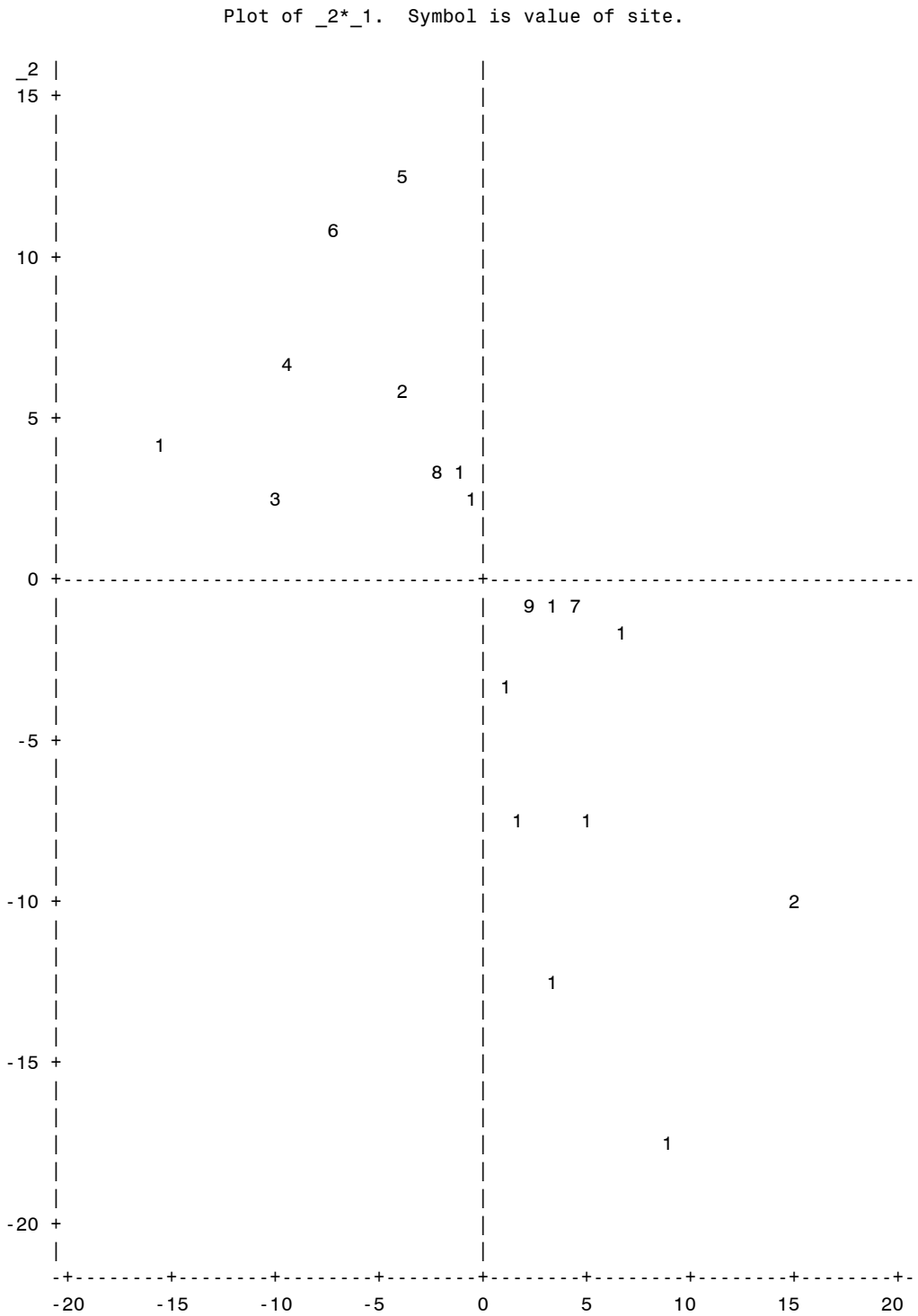


Figure 1. Residual of treatment 2 vs. treatment 1, Example 1.

One observation (site 1) has the missing observation and is not plotted. The graph shows that treatment 1 and 2 residuals cluster by site in either the upper left quadrant (high treatment 2 residual) or lower right quadrant. The graph suggests that some sites have an equal and opposite treatment effect. Going back to the sites and looking at the procedures for assigning treatment, suppose it was found that sites 11 through 20 had somehow reversed the treatment assignment. After determining this, an analysis is done to the data using the actual treatments applied, and a simpler picture appears: the re-run shows that the simple treatment effect is significant, with no interaction!

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	1	1557.890890	1557.890890	18.51	<.0001
site	18	783.354291	43.519683	0.52	0.9366
site*trt	18	1404.776637	78.043146	0.93	0.5519

Suppose that no problem in treatment assignment was found, however. In this case, treatment effect can be found within the site clusters. Using the variable SiteGrp to distinguish between the two clusters, each cluster is shown to have significant treatment effect:

```
proc glm data=clin3; by SiteGrp;
title Analysis Showing Significant Interaction;
class site trt;
model y=trt|site;
estimate 'Treatment' trt -1 1;
run;
```

The simple treatment effect within sites 1-10 is not estimable because only one treatment arm is available for site 1. When this site is removed, Sites 2-10 show approximately equal and opposite treatment effect from sites 11-20.

The next step would be to look for patient or site characteristics that would actually reverse the effect for half of the sites.

Example 2: Effect Variation by Site

It would be an important finding if site by treatment interaction existed and that it was inherent in the type of subjects treated at different sites. For example, what if half the sites had twice the treatment effect as the other half?

GLMPOWER[3] is a procedure now available in version 9 SAS®, and can be used to plan for sample size when considering interaction effects. GLMPOWER uses as input an "exemplary dataset", which embodies the design and response variables. Our exemplary dataset uses the design of the previous example, with two levels of treatment, 20 sites, an effect size of 10 for sites 1-10, but doubling the effect size for sites 11-20.

GLMPOWER is run in two scenarios, one using site as a class variable, and one using "Ovrsite", an alias for whether site is greater than or less than 10. The code is as follows:

```
data dsite;
do site=1 to 20; ovrsite= int((site-1)/10);
eff=10;
do trt=1 to 2;
if trt=1 then eff=0;else if trt=2 then eff=10;
if site>10 then eff=eff*2;   *** Double effect at site >10;
output;
end;
end;
run;

****************** Scenario 1:  Site effect and interaction ******************;
proc glmpower data=dsite;
class site trt;
model eff=site|trt;
power alpha=.05 stddev=10 ntotal=100 power=.;
```

Scenario 1: Site as Class variable	
Dependent Variable	eff
Alpha	0.05
Error Standard Deviation	10
Nominal Total Sample Size	100
Actual Total Sample Size	80
Error Degrees of Freedom	40

Source	Test DF	Power
site	19	0.149
trt	1	>.999
site*trt	19	0.149

Using this scenario, both site and site*trt, with 19 degrees of freedom, has less than 15 % chance of being found significant, for alpha = 0.05. A much larger sample size would be required in order to have a good power of getting a significant result.

Under scenario 2, we repeat the analysis but using the alias “OvrSite” variable, we see that chances are much better of being able to see a significant site and site*trt effect, nearly 60% power. To run scenario 2, the following statements are used:

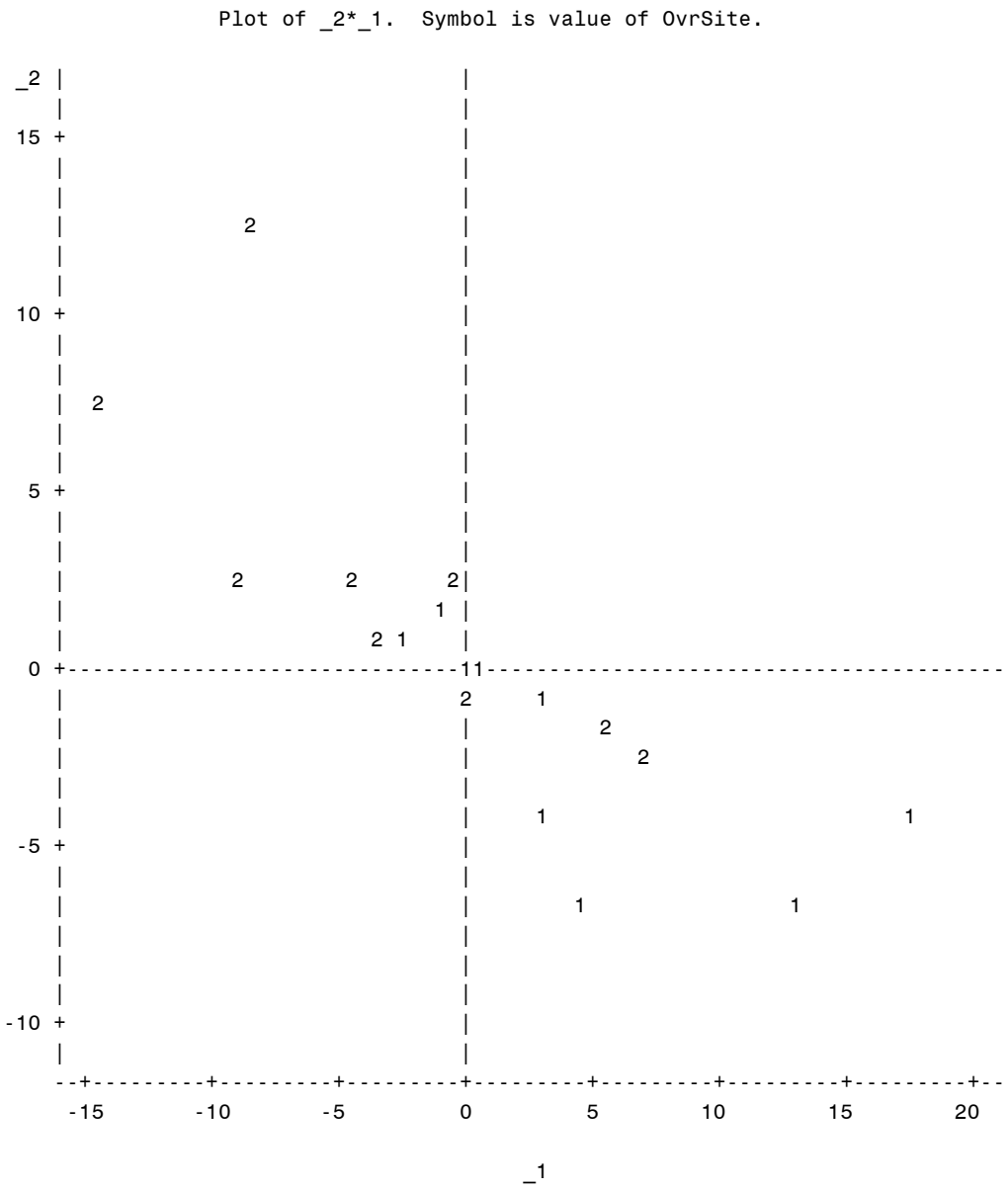
```
proc glmpower;
class ovrsite trt;
model eff=ovrsite|trt;
power alpha=.05 stddev=10 ntotal=100 power=.;
run;
```

Scenario 2: OvrSite as Class variable	
Dependent Variable	eff
Alpha	0.05
Error Standard Deviation	10
Nominal Total Sample Size	100
Actual Total Sample Size	80
Error Degrees of Freedom	76

Source	Test DF	Power
ovrsite	1	0.598
trt	1	>.999
ovrsite*trt	1	0.598

Using Ovrsite, with 1 df in the model, instead of site, with 19 df increases power. By looking at the residual plot and looking for common factors of sites that cluster together, one might find covariates which would explain the clusters. The sample size must suffice to allow the mean residual values to be located with enough precision to distinguish clusters. However since the power to detect the treatment effect is high, clusters may be readily distinguishable.

Using the previous dataset with correct treatments applied, effect variation by site is introduced, with an effect size of 10 for Treatment 2 when Ovrsite=1 and 20 for Ovrsite=2, adjusting for main effects and plotting the residual we have the next Figure:



NOTE: 1 obs had missing values. 1 obs hidden.

Figure 2. Residual of treatment 2 vs. treatment 1, Example 2.

Again, the residuals of the main effect model plot the points in the upper left and lower right quadrant. For OvrSite = 2, treatment 1 is over-compensated and treatment 2 is under-compensated, plotting them in the upper left quadrant. Likewise, OvrSite = 1, treatment 1 is under-compensated and treatment 2 is over-compensated, plotting them in the lower right quadrant.

Discussion

Site Pooling is sometimes offered as a way to allow for estimation of site-by-treatment interaction. When sites have few patients per treatment group, these small sites are pooled until a certain minimum number of subjects per treatment-site combination are achieved.

Pooling sites solely on the basis of low enrollment however does not seem likely to uncover any interesting site-by-treatment interaction, and it would obscure any problems caused by treatment assignment or other artifactual issues discussed previously.

A more promising approach is to look for possible confounding factors, variables which modify the treatment effect, and, where possible, assure balanced strata across sites, eg, if patient age is a confounder, enroll a uniform distribution of subjects across sites. In addition, just because sites are uniform in their distribution of subjects, there may still be between-subject, within site variability that can be controlled, so that the variable should be used as a covariate in the analysis, including it in an interaction term with treatment. If interaction is ruled out, or accounted for by covariates, the site-by treatment interaction term may be dropped from the model.

Conclusions

Analysing and plotting mean residuals may provide clues to explain site by treatment interaction. Examples using two treatments and twenty sites have been shown. With three or more treatments, a cluster analysis approach could be used to look for clusters of sites. Because of the ad-hoc nature of these analyses, findings which result should be followed up for confirmation with independent studies.

REFERENCES

1. Guidance for Industry E9 Statistical Principles for Clinical Trials, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), September 1998

2. Winer, B. J., Statistical Principles in Experimental Design (Second Edition), McGraw-Hill Book Company, 1971, pp. 318-319.

3. PROC GLMPOWER: SAS/STAT(R) 9.2 User's Guide, Copyright © 2009 SAS Institute Inc. All Rights Reserved.

ACKNOWLEDGMENTS

I would like to thank Mary Poole of Pharmanet for encouraging my work on this topic.

CONTACT INFORMATION

Your comments and questions are valued and encouraged:

Alan Davis, PhD
PharmaNet, Inc.
10150 Mallard Creek Rd.
Suite 500
Charlotte, NC 28215
(704)602-7904
adavis@pharmanet.com