Paper 195-31

Using SAS® To Investigate Effect Modification

Peter H. Van Ness, Heather G. Allore
Program on Aging, Department of Internal Medicine,
Yale University School of Medicine, New Haven, CT

ABSTRACT

Effect modification occurs in biomedical research when a measure of statistical association between an exposure and a health outcome, as represented on some specific scale, differs according to the levels of a third variable—the effect modifier. The heterogeneity of a measure of association across the levels of a modifying variable has an importance that might be statistical, epidemiologic, and/ or biologic. Respectively, it might contribute to the goodness of fit of a regression model, identify subgroups for which a health risk factor is especially prominent, and/ or specify a multifactorial causal relationship. Effect modification can be tested and graphed in numerous ways with SAS statistical software. In this article we will provide code for six different ways of investigating effect modification, primarily using PROC GENMOD, and two ways of graphing statistical interactions using the GPLOT and (via ActiveX Control) the G3D procedures. Advantages and disadvantages of each approach will be stated.

Statistical interaction should not be confused with biological interaction. Causal biologic relationships can only be analyzed meaningfully in longitudinal studies. This article will conclude with a discussion of the relationship between representing statistical interactions in SAS software and drawing causal statistical inferences in longitudinal studies. Specifically, the strengths and limitations of fitting marginal structural models for time-varying confounding with PROC GENMOD will be identified.

INTRODUCTION TO EFFECT MODIFICATION

Upon regressing an outcome variable on an exposure variable, epidemiologists characteristically next investigate potential confounding factors and effect modifiers (Harrell 2001). Confounders—variables associated with both the exposure and outcome that alter the association between them but are not in their causal pathways—are not of substantive interest. Their inclusion in a regression model is intended to prevent bias in estimates of a hypothesized association. Contrarily, effect modifiers are often of substantive importance. In some circumstances interaction terms might be included in regression models solely to promote goodness-of-fit, i.e., their interpretation as specification of subgroup associations or biologic interactions is implausible but their presence in the model promotes a distribution of residuals that better satisfies model assumptions. In most applications, however, statistically significant interactions between covariates and a main predictor are qualifications or extensions of hypothesized associations that merit careful representation in a report of a biomedical research study. Hence, identifying different ways in which effect modification can be investigated with SAS software, and specifically, with PROC GENMOD, is not only a matter of technical practice but also of inferential interpretation.

Effect modification occurs when a measure of statistical association between an exposure and an outcome, as represented on some specific scale, differs according to the levels of a third variable—the effect modifier. The statistical significance of an interaction term crossing an exposure predictor with a potential modifier establishes the modification of the measure of association and provides evidence for a modification of a sociologic or biologic effect. (For simplicity we will speak henceforth of “predictors” and “modifiers” recognizing that the model might include more predictors than the one crossed with a modifier and that the modifier is only potential until shown to be statistically significant.) Clearly, presence of statistical interaction is not sufficient to establish the presence of causal biologic interaction. To emphasize this point Greenland and Rothman prefer the term “effect measure modification” (Greenland and Rothman 1998). Last makes the same point differently by defining an “effect modifier” as the modification of “the effect of a *putative* causal factor” (emphasis added) (Last 2001). Mindful of these important interpretive caveats, this article uses the variant of the term common in epidemiology and biostatistics (Miettinen 1974).

Six Coding Approaches

A strength of SAS statistical software is its flexibility and comprehensiveness. Accordingly, there are numerous ways that an analyst can investigate suspected effect modification. Six will be described here; code for each is provided in Appendix A. For clarity we will consider the case in which the outcome, predictor, and modifier variables are all binary.

The Analytic Approach: The simplest approach uses the FREQ procedure to generate two 2 x 2 tables, one for each level of the modifier (Appendix A.1). The extent to which the odds ratios for the two tables differ is the extent to which modification occurs. A Breslow-Day statistic—generated by adding the "chm" option to the "tables" statement—tests the homogeneity of the odds ratios across the levels of the modifier. The advantage of this approach is that it allows you to inspect the joint distributions of the two tables, identifying low cell counts or complete separations if they occur. Also, with appropriate coding of the variables, it allows for easy calculation and comparison of odds ratios and risk ratios as measures of association. Its disadvantage is that it doesn't explicitly provide all potentially relevant information, e.g., residuals and other diagnostic information, and even p-values for the odds ratios and risk ratios. Of course, this approach is not appropriate when the outcome or predictor is measured on a continuous or count scale or when there are many additional covariates.

The Stratified Regression Approach: A second approach is to perform stratified logistic regressions using the "by" statement to generate separate analyses for each level of a modifier (Appendix A.2). This approach yields a wealth of diagnostic and output information, especially when the "influence" option is used in the LOGISTIC procedure and the "obstats" option is used with PROC GENMOD. It easily handles models with continuous and count variables and many covariates. Perhaps its most attractive feature is that it allows for a very intuitive understanding of effect modification. If a more complicated coding strategy is used—such as those to be discussed below—and one wants to confirm an interpretation of modifier level specific results, then comparing these results with those from stratified models serves that purpose well.

This stratified approach does not provide a test of statistical significance of the difference between the stratified odds ratios. A more serious statistical shortcoming arises when a model has numerous covariates in addition to the modifier. Stratification unnecessarily attenuates multicollinearity among the covariates because it allows for no statistical interrelationships between data items segregated into the stratified models. One might think that because interaction terms are highly correlated with their crossed components that stratification would be recommended to prevent bias from this sort of collinearity. It will be discussed later how centering interaction terms provides a better solution to this problem. Centering and other subsequent approaches allow for complicated interrelationships among covariates, and although the simple models provided in the appendix all yield the same parameter estimates this would not be the case for models with many additional covariates. The stratified models would provide slightly different and less satisfactory results. Thus, interpretations of measures of association for stratified models are also subtly different: statistical inferences can be generalized only to the population from which the sample stratum was drawn and not to the entire original sample.

The Interaction Term Approach: The third approach to investigating effect modification involves crossing a predictor and modifier and adding this interaction term to the regression model (Appendix A.3). This approach is the one most commonly presented in introductory regression textbooks (Kleinbaum et al. 1998). When the modifier is binary and 0/1 coding is used to indicate respectively the absence and presence of some characteristic, then the parameter estimate and standard error of the main predictor are interpretable as their values for the 0 level of the modifier. If one runs a second model using inverted coding for the modifier, then the results provide information comparable to the 1 level of the modifier. The greater the number of levels for the modifier, the more burdensome it is to run separate models in order to obtain all desired information.

The Specified Levels Approach: The shortcoming just described for the interaction term approach can be overcome by a slight modification that omits the main predictor of interest from the model and includes not only an interaction of this main predictor and the modifier but also an interaction of this main predictor and an inversely coded version of the modifier (Appendix A.4). This technique cannot be used when the modifier has more than two levels. The fact that the interaction term approach to investigating effect modification requires several models to obtain all requisite information is related to the lack of interpretability of its parameter estimates in isolation. Use of 0/1 coding partially overcomes this lack of interpretability but is not always a viable option, e.g., when discrete modifier levels result from categorization of a continuous variable and require precise numeric specification. Beginning analysts are frequently cautioned not to interpret main predictors and interaction terms in isolation from one another. This is generally sound advice

but does not apply to all model parameterizations for investigating effect modification. For instance, one exception is a model that has only a modifier variable and a variable crossing the modifier with a main predictor: the predictor does not occur alone in the model. In this case, the parameter estimate of the interaction term is, in fact, the parameter estimate of the main predictor for persons with a value of 1 on the modifier variable. If the "class" statement is used along with this parameterization then parameter estimates for the association of the predictor and outcome are yielded for both levels of the modifier (See Appendix A.5.2).

The Nesting Approach: This model is equivalent to a model in which the modifier occurs alone and the predictor is "nested" within the modifier (Appendix A.5.1). In a way this is not surprising because the design matrix for nested variables is the same as the design matrix for crossed variables. The most detailed discussion of nested and crossed effects in the *SAS/ STAT User's Guide* occurs in the material on the MIXED procedure subtitled "Parameterization of Mixed Models" (*SAS/ STAT User's Guide* 2003). It notes that nested variables approximate "random effects within a fixed effects framework." Furthermore, it states that with the advent of the "random" statement in SAS mixed model procedures the specification of nested effects in the model statement becomes unnecessary. What remains important about the nested approach to investigating effect modification is its interpretive perspective. Effect modification entails the specification of measures of predictor/ outcome association for all levels of a modifier variable. Not only is the predictor said to be nested within the modifier but both these variables and their interaction constitute a regression model that is itself nested as a special case within the model consisting only of the predictor and modifier independent variables (Agresti 2002). As such, the two models satisfy the condition for a likelihood ratio test, and, in fact, the difference in deviances between the two models provides the value for the likelihood ratio test of the interaction term as provided by the "type3" model option. The notion of stratification suggests division; nesting suggests integration. Only in relatively simple—and, so relatively abstract—models can investigating effect modification be identified with the repeated calculation of measures of association on segregated strata of data. More generally, investigating effect modification is part of a complex process—fitting, testing, and refitting--of identifying a model that best combines simplicity of representation with adequacy to the data represented.

The Centered Interaction Term Approach: A final approach to investigating effect modification recommends the centering of modifier and predictor variables and their interaction term (Appendix A.6.). Advocates of this approach say that it guards against collinearity among this set of variables and provides intelligible meanings to them both individually and in combination (Kraemer and Basey 2004). Certainly, centered versions of variables used in investigating effect modification—usually coded -0.5 and 0.5 for a binary modifier—show less correlation than their noncentered counterparts. However, the degree of correlation in the noncentered variables is not usually so great as to introduce bias in standard errors or parameter estimates. The example in Appendix A demonstrates this.

The interaction term in the centered approach has the same value as the interaction term in the noncentered interaction term approach. It represents the difference in parameter estimates for the predictor-outcome association for the two levels of the modifier. Its exponentiation is not interpretable. The parameter estimate for the centered predictor represents the unweighted average of the parameter estimates for the predictor-outcome association for the two levels of the modifier variable. Although interpretable, this average parameter estimate is abstract, corresponding to no actual group of study participants. For instance, it is not generally equal to the parameter estimate of the predictor-outcome association in the absence of an interaction of the predictor with the modifier. Thus, the claimed advantage in interpretability is somewhat specious.

The parameter estimates yielded by the centered approach can be used to calculate the parameter estimates for the association of interest for the two levels of the modifier. Such calculations are somewhat more complicated for this approach than for the others mentioned. One subtracts half of the parameter estimate for the interaction term from the parameter estimate for the centered predictor and this yields the parameter estimate for the 0 level of the modifier. Adding in the above calculation will yield the parameter estimate for the 1 level of the modifier.

Summary of the Six Approaches: Choice among the six approaches for investigating effect modification should be sensitive to the circumstances of the data analysis. If the data is categorical examining the joint distribution of predictor and outcome variables across the levels of the modifier is informative. If collinearity is a concern then the centered approach is recommended. In some cases the study design might dictate use of stratified regression models, e.g., when random sampling of study participants occurs within strata. There is probably no circumstance where the nested approach is now required. The interaction term approach is probably most widely applicable.

TWO GRAPHING APPROACHES

In most statistical analyses it is helpful to examine appropriate graphical representations of the data. Graphical presentations of effect modification can sometimes be quite dramatic.

PROC GPLOT: Figure 1 was generated by PROC GPLOT (graphs and code presented in Appendix B) and indicates the presence of a qualitative interaction, i.e., one in which the direction as well as the magnitude of the measure of association changes across the levels of the modifier. The lines cross in this case; in quantitative interactions the lines do not cross but have slopes that differ in statistically significant ways. Here predicted probabilities are graphed for both levels of the modifier across the levels of the predictor variable. Actually, the measure of association represented is the risk ratio rather than the odds ratio. PROC GENMOD does not automatically generate the predicted odds for unique combinations of predictor variable levels—predictor profiles—but they could be calculated from the predicted probabilities with a little programming.

PROC G3D: Since effect modification involves the relationship between three variables it is conducive to being plotted in three-dimensional space. PROC G3D will accomplish this. In fact, the procedure provides a more flexible implementation of a scatter plot matrix for three variables—a graphical innovation that emerged from Bell Labs in the early 1980's (Chambers et al 1983). Such a plot is of little utility for the categorical data used in Appendix A, so continuous variables are used for the outcome and predictor in Figures 2 and 3 in Appendix B. If the modifier is also continuous, it is advisable to categorize it into a small number of discrete levels to promote insightful visual interpretation. Also, not all perspectives on the three-dimensional configuration of data points will be revealing. Figure 1 reveals little structure relevant to effect modification. However, by right clicking on the three-dimensional graph created by the given code a menu will appear that identifies the "Graph Toolbar." By clicking on this menu item and then by clicking on the "Rotate" button, the analyst can use the left mouse button to grab the edge of the three-dimensional plot and rotate it in a way that is more insightful. Figure 3 is one such rotation that shows the respectively ascending and descending pattern of outcome values over the values of the predictor variable for the 0 and 2 levels of the modifier. Three-dimensional scatter plots are helpful for an initial exploration of effect modification; a two-dimensional plot can instructively confirm its presence.

A CONCLUDING NOTE ON LONGITUDINAL STUDIES

A primary purpose of longitudinal studies is to show the change over time of the mean value of an outcome. This can be accomplished by including in a longitudinal regression model an interaction between a time variable, often time from baseline or age, and the main predictor of interest (Fitzmaurice et al 2004). This is a special case of effect modification. It can occur in combination with effect modifications discussed previously. Especially relevant are biologic interactions that can most plausibly be established in a longitudinal context. However, special types of model fitting challenges arise when trying to draw causal inferences. For instance, time-varying confounders can occur, i.e., a confounder of a hypothesized causal effect can also be a mediator of this relationship. In an epidemiologic study of a medication for some diseases, it is possible that participants will increase the dosage of their medication if their illness symptoms worsen or fail to improve. Such updated medication exposures are potential time-varying confounders of the original exposure-outcome association. New models have been proposed to account for this sort of confounder and insure valid causal inference. Robins has proposed "structural nested" and "marginal structural" models for this purpose (Robins et al. 1999). The first type of model allows for the investigation of effect modification using the interaction term approach but cannot be fit with SAS software. The second type of model can be fit with PROC GENMOD but does not permit examination of time-varying interactions (Robins et al. 2000). Thus knowing various approaches for investigating effect modification with the SAS system enables a data analyst to both enjoy its full flexibility and respect its current limitations.

ACKNOWLEDGEMENTS

This work was supported in part by NIH Grant # P30AG21342. The authors thank Zhenchao Guo for his assistance with the three-dimensional graph.

APPENDIXES

```
*Appendix A: Six Ways of Numerically Investigating Effect Modification;
data modification;
      input  modifier predictor outcome count;
inv_modifier  = abs(modifier - 1);      *Inverted coding of variables;
inv_predictor = abs(predictor - 1);
inv_outcome   = abs(outcome - 1);
cen_modifier  = modifier - 0.5;  *Centered coding of variables;
cen_predictor = predictor - 0.5;
cen_outcome   = outcome - 0.5;
datalines;
0      0      0      46
0      0      1      75
0      1      0      11
0      1      1      54
1      0      0      11
1      0      1      67
1      1      0      17
1      1      1      44
;
run;

*1    The Analytic Approach;
proc freq data=modification;
      weight count;
      tables modifier*inv_predictor*inv_outcome / measures cmh;
run;
*Note that SAS compares lower levels to higher levels of variables in PROC FREQ
so variable coding must be inverted in order to get appropriate relative risk
results.  The odds ratio and most other measures of association produced by
this procedure are invariant to the coding transformation;
*2    The Stratified Regression Approach;
proc sort data=modification;
      by modifier;
run;
proc genmod data=modification descending;
      freq  count;
      by modifier;
      model outcome = predictor / dist=binomial link=logit /*log*/ /*obstats*/;
      estimate 'Main Predictor OR Calculation'  predictor 1 / exp;
run;
*Note that by changing to the log link the regression model produces risk
ratios instead of odds ratios;
*3  The Interaction Term Approach;
proc genmod data=modification descending;
      freq  count;
      model outcome = predictor modifier modifier*predictor
/*inv_modifier*predictor*/ / dist=binomial link=logit type3;
      estimate 'Modifier Absent /*Present*/ OR Calculation'
                  predictor 1 / exp;
run;
*4    The Specified Levels Approach;
proc genmod data=modification descending;
      freq  count;
      model outcome = modifier inv_modifier*predictor modifier*predictor /
dist=binomial link=logit type3;
      estimate 'Modifier Absent OR Calculation'
            inv_modifier*predictor 1 / exp;
      estimate 'Modifier Present OR Calculation'
            modifier*predictor 1 / exp;
run;
```

```
*5  The Nested Approach;
proc genmod data=modification descending;
       freq  count;
       class /*inv_*/modifier;
       model outcome = /*inv_*/modifier predictor(/*inv_*/modifier)  /
       dist=binomial link=logit type3;
       estimate 'Modifier Absent /*Present*/ OR Calculation'
       predictor(/*inv_*/modifier) 1 / exp;
run;
*Note that the nested approach is equivalent to the interaction term approach
from which the predictor has been removed as a main effect;
proc genmod data=modification descending;
       freq  count;
       class modifier /* / ref=first param=ref */;
       model outcome = modifier predictor*modifier  / dist=binomial link=logit
       type3;
       estimate 'Modifier Absent /*Present*/ OR Calculation'
       predictor*modifier 1 / exp;
run;
*6  The Centered Interaction Term Approach;
proc genmod data=modification descending;
       freq  count;
       model cen_outcome = cen_modifier cen_predictor cen_modifier*cen_predictor
/ dist=binomial link=logit type3;
       estimate 'Main Predictor OR Calculation'  cen_predictor 1 / exp;
run;
*Note that the parameter estimate (log odds ratio) in the above centered model
is the unweighted average of the parameter estimates for the predictors in the
two models stratified by the moderator;
*Note that the odds ratios for the association of the predictor to the outcome
for the individual levels of the modifier is obtained by adding and subtracting
half of the value of the interaction term parameter estimate to the parameter
estimate of the predictor and then exponentiating the results.

*Appendix B: Two Ways of Graphically Representing Effect Modification;

*1     The Two-Dimensional Approach Using PROC GPLOT;
proc genmod data=modification descending;
       freq count;
       model  outcome = predictor modifier modifier*predictor / dist=bin
       link=logit;
       output out=output pred=pred;
run;
goptions reset=global ftitle=swissb ftext=swissi htitle=3 htext=2
/*device=activex*/;
title 'Interaction Plots';
symbol1 i=join v=dot     line=2 width=2 color=black;
symbol2 i=join v=diamond line=1 width=3 color=black;
axis1 order = (0 to 1 by 1)
         label = ('Predictor Values')
         major= ( width=2) minor= (width=2)
         width=3;
axis2 order = (0 to 1 by 0.1)
         major=(width=2) minor = (width=2)
         label = (a=90 'Predicted Probabilities')
         width=3;
legend1       label=('Modifier Levels:')
                    position=(bottom center inside)
                    mode=share;
proc gplot data=output;
       plot   pred*predictor=modifier / legend=legend1 haxis=axis1
                                        vaxis=axis2;
run; quit;
```

```
*2     The Three-Dimensional Approach Using PROC G3D;
goptions device = activex; *This tells SAS to create an ActiveX object;
ods listing close;         *This closes the listing output so you don't
                           get the graph in 2 places;
ods html path='p:/van_ness/poa articles/modification/programs'
body='plot1.html' (url=none);
                              *This creates an html file that can be
                              opened in Internet Explorer;
data modification_continuous;
       input modifier predictor outcome @@;
datalines;
0      1      5      1      1      11     2      1      19
0      2      7      1      2       9     2      2      18
0      3      9      1      3       8     2      3      15
0      4      11     1      4       7     2      4      14
0      5      12     1      5       9     2      5      14
0      6      17     1      6       8     2      6      18
0      7      16     1      7       6     2      7      12
0      8      17     1      8       7     2      8       9
0      9      19     1      9       9     2      9       7
0      10     20     1      10     10     2      10      5
;
run;
/* Axes Statements */
axis1
    label=('modifier')
    minor=none;
axis2
    label=('predictor')
    minor=none;
axis3
    label=('outcome')
    minor=none;
title  'Three-Dimensional Scatter Plot:Original /*Rotated*/ View';
proc g3d data=modification_continuous;
       scatter modifier*predictor=outcome /
    zaxis=axis1
    xaxis=axis2
    yaxis=axis3
    grid
    noneedle
    size=2
       color='Red';
run;
quit;
goptions reset=all;
ods html close;            *This closes the html output destination;
ods listing;               *This reopens the output window in SAS;
```

FIGURES

Figure 1

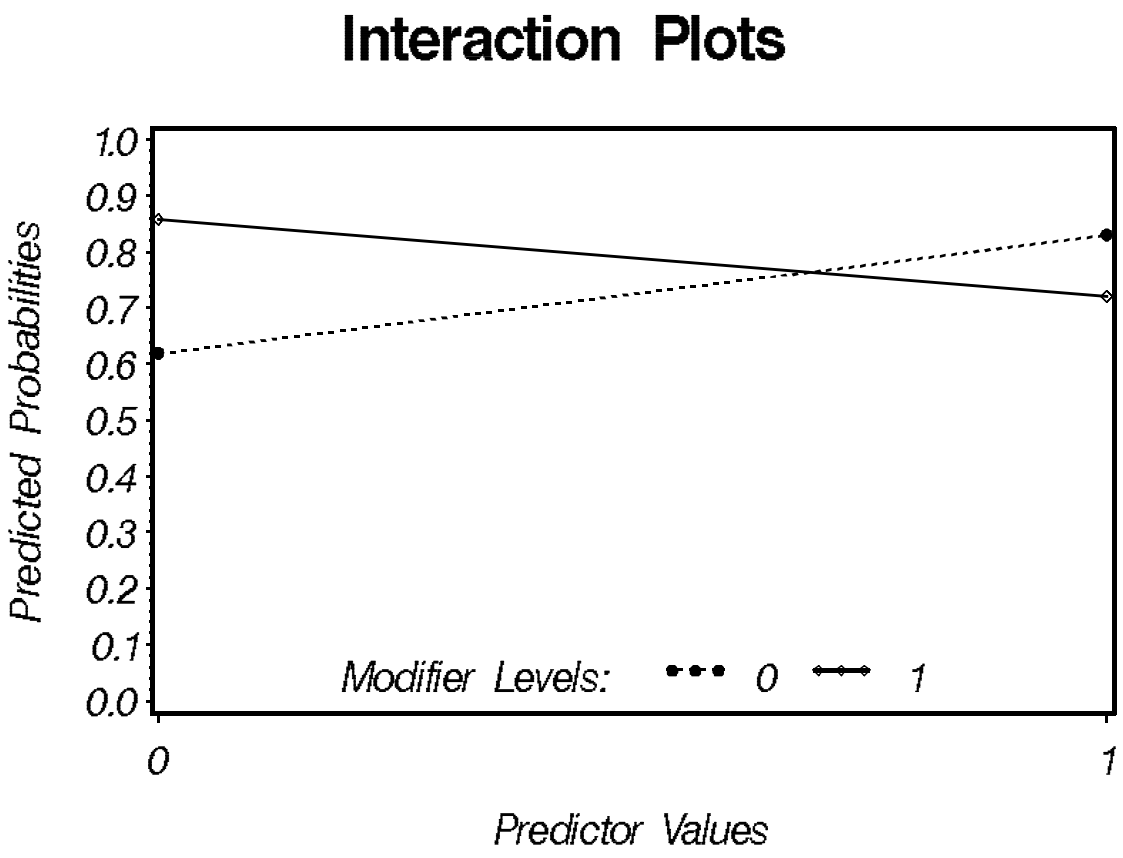


Figure 2: Three-Dimensional Interaction Scatter Plot: Original View

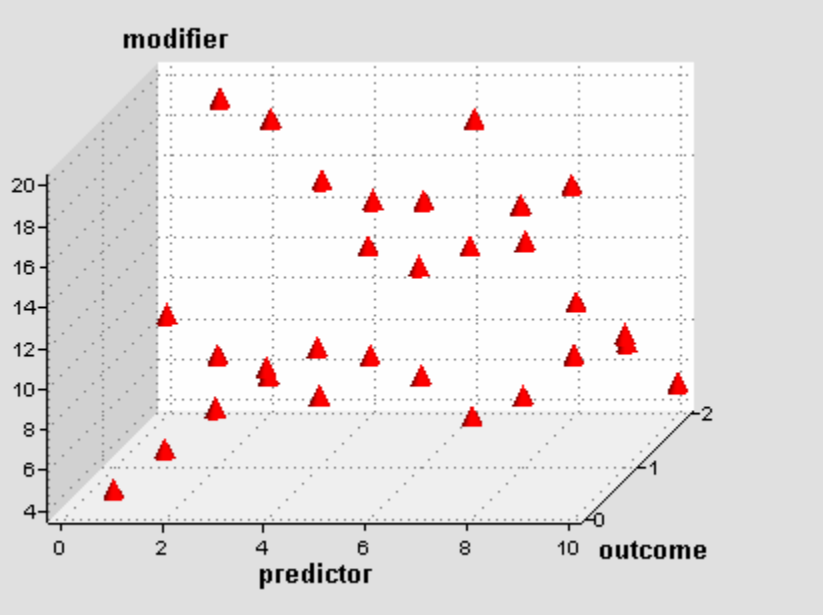
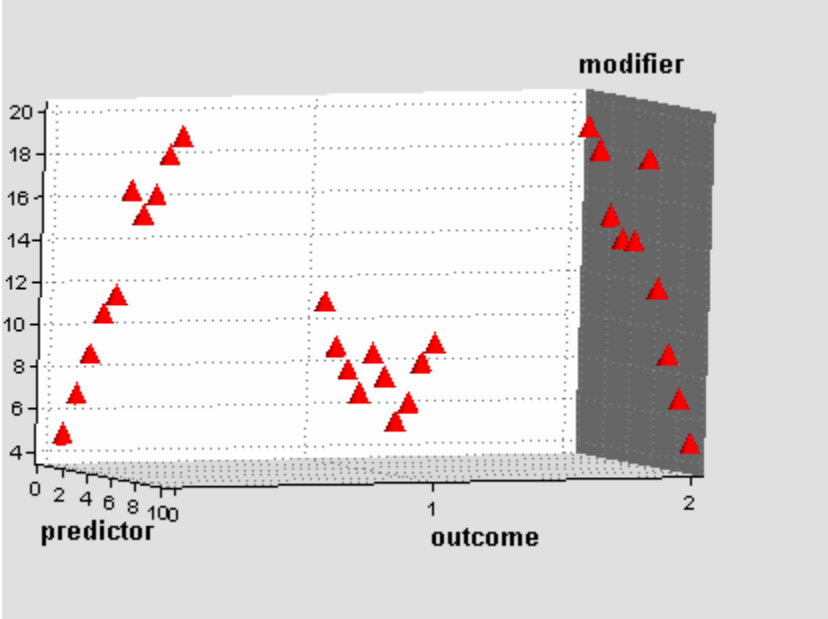


Figure 3: Three-Dimensional Interaction Scatter Plot: Rotated View



REFERENCES

Harrell FE, Jr. 2001. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer.

Greenland S, and Rothman KJ. 1998. "Concepts of Interaction." In *Modern Epidemiology*, edited by Kenneth J. Rothman and Sander Greenland, 329-42. New York: Lippincott-Raven.

Last JM, ed. 2001. *A Dictionary of Epidemiology*. 4th ed. New York: Oxford University Press.

Miettinen O. 1974. Confounding and effect modification. *American Journal of Epidemiology* 100:350-53.

Kleinbaum DG, L. KL, Muller KE, and Nizam A. 1998. *Applied Regression Analysis and Other Multivariable Methods*. 3rd ed. Pacific Grove, CA: Duxbury Press.

***SAS/ STAT User's Guide*. 2003. Cary, NC: SAS Institute, Inc.**

Agresti A. 2002. *Categorical Data Analysis*. 2nd ed. Hoboken, NJ: John Wiley & Sons.

Kraemer HC, and Basey CM. 2004. Centering in regression analysis: a strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research* 13:141-51.

Chambers JM, Cleveland WS, Kleiner B, and Tukey PA. 1983. *Graphical Methods for Data Analysis*. Boston: Duxbury Press.

Fitzmaurice GM, Laird NM, and Ware JH. 2004. *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons.

Robins JM, Greenland S, and Hu F-C. 1999. Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome (with discussion). *Journal of the American Statistical Association* 90:106-21.

Robins JM, Hernan MA, and Brumback B. 2000. Marginal structural models and causal inference in epidemiology. *Epidemiology* 11:550-60.

CONTACT INFORMATION

Name: Peter H. Van Ness
Enterpreise: Yale University School of Medicine
Address: Program on Aging, 1 Church Street, 7th Floor
New Haven, CT 06510
Work Phone: 203-764-9886
Work Fax: 203-764-9831
E-mail: peter.vanness@yale.edu
Web: http://www.info.med.yale.edu/intmed/geriatrics/pages_faculty/vanness.htm