ODDS RATIOS IN A TABULAR PRESENTATION
Stephen M. Noga, UNC-CH, Chapel Hill, NC
Ding Yi Zhao, UNC-CH, Chapel Hill, NC

Abstract
Odds ratios and their corresponding confidence intervals are easily obtainable statistics using the 'RL' option in PROC LOGISTIC (SAS®/STAT). However, there is no option to output these statistics to a SAS dataset. This presents a challenge to the user when the task at hand is to present those statistics in tabular form, such as output from PROC TABULATE. This paper uses PROC LOGISTIC, PROC TABULATE, and the SAS Macro language to produce the desired tables.

Introduction
So you want to create a table of odds ratios and their 95% confidence intervals, but you only want some of the independent variables listed in that table, and you think you might like to see the results from a number of different models adjacent to each other in this table. On the surface, this does not seem to be an overwhelming problem. After all, that is what PROC TABULATE is for. The $1.98 question is: "How do I get the necessary statistics from PROC LOGISTIC into PROC TABULATE?".

The SAS code to answer that question follows along with a running commentary. This paper is intended for all levels of SAS users. The code was written in version 6.10 of SAS using the OS/2 operating system, but it is not system dependent.

Onward
Let us first look at the output created by using the 'RL' option in the PROC LOGISTIC model statement.

PROC logistic data=man.uc24;
  model MMNSA_OX = PD PH PDH PDD PHH PDDH PHHD PDDHH VIAGEO1 SEX D1 D2 D3 D4 / rl;
RUN;

This produces the following output:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Odds Ratio</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>1.0000</td>
<td>1.536</td>
<td>1.013</td>
<td>2.329</td>
</tr>
<tr>
<td>PH</td>
<td>1.0000</td>
<td>1.422</td>
<td>1.058</td>
<td>1.911</td>
</tr>
<tr>
<td>PDH</td>
<td>1.0000</td>
<td>2.481</td>
<td>1.743</td>
<td>3.530</td>
</tr>
<tr>
<td>PDD</td>
<td>1.0000</td>
<td>3.831</td>
<td>1.304</td>
<td>11.258</td>
</tr>
<tr>
<td>PHH</td>
<td>1.0000</td>
<td>2.280</td>
<td>1.484</td>
<td>3.501</td>
</tr>
<tr>
<td>PDDH</td>
<td>1.0000</td>
<td>6.357</td>
<td>2.813</td>
<td>14.362</td>
</tr>
<tr>
<td>PHDH</td>
<td>1.0000</td>
<td>4.446</td>
<td>2.820</td>
<td>7.009</td>
</tr>
<tr>
<td>PDDHH</td>
<td>1.0000</td>
<td>7.134</td>
<td>2.661</td>
<td>19.123</td>
</tr>
<tr>
<td>VIAGEO1</td>
<td>1.0000</td>
<td>1.128</td>
<td>1.106</td>
<td>1.152</td>
</tr>
<tr>
<td>SEX</td>
<td>1.0000</td>
<td>1.035</td>
<td>0.826</td>
<td>1.296</td>
</tr>
<tr>
<td>D1</td>
<td>1.0000</td>
<td>0.947</td>
<td>0.661</td>
<td>1.356</td>
</tr>
<tr>
<td>D2</td>
<td>1.0000</td>
<td>3.463</td>
<td>2.498</td>
<td>4.802</td>
</tr>
<tr>
<td>D3</td>
<td>1.0000</td>
<td>1.592</td>
<td>1.151</td>
<td>2.202</td>
</tr>
<tr>
<td>D4</td>
<td>1.0000</td>
<td>2.705</td>
<td>1.476</td>
<td>4.956</td>
</tr>
</tbody>
</table>

This output has everything you need, but your degree of satisfaction is low because the default output from PROC LOGISTIC is not quite what you need for your presentation.

On the following pages, you will see code for the macro which is used for all calculations and code that calls the macro and produces the tables. The first table simply replicates the PROC LOGISTIC output using PROC TABULATE to pretty it up. The second and third tables expand on this by having multiple models across the table and multiple categories down the table.
The Macros

Title: LOGISTIC

Description: This program uses SAS PROC LOGISTIC to calculate odds ratios and 95% confidence limits for each explanatory variable.

Programmer: Dingyi Zhao
Modified By: Steve Noga

Department of Biostatistics
University of North Carolina at Chapel Hill

You have to provide the following macro variables.

1. %let x_num = K how many explanatory variables
2. %let indata = analysis data set.
3. Zlet var_y = Y response variable(s).
4. %let var_x = X1 X2 ... Xk K explanatory variable(s).
5. %let drpvar = variables to be dropped

The following is the process for calculating odds ratios and 95% confidence limits.

The first, using PROC LOGISTIC with OUTEST and COVOUT options.

PROC LOGISTIC DATA = &INDATA OUTEST = TEMP COVOUT
MODEL &VAR_Y &VAR_X;

AFTER RUNNING, ONE GETS VARIANCE AND COVARIANCE MATRIX.

_NAME_ INTERCPT X1 X2 ...... Xk
ESTIMATE B1 B2 ...... Bk
X1 COV11 COV12 ...... COV1k
X2 COV21 COV22 ...... COV2k
......
Xk COVk1 COVk2 ...... COVkk

LET B(i) = Bi i = 1,2, ... , k
VAR(i) = COVii I = 1,2, ... , k

THEN ODDS RATIOS FOR EACH EXPLANATORY VARIABLE IS
OR(i) = EXP(B(i)) i = 1,2, ... , k

95% CONFIDENCE LIMITS
L(i) = EXP(B(i) - 1.96*SQRT(VAR(i)))
U(i) = EXP(B(i) + 1.96*SQRT(VAR(i)))

WALD STATISTIC FOR EACH EXPLANATORY VARIABLE X(i) IS
WALD = (B(i)*B(i))/VAR(i). CHI SQUARE WITH 1 DEGREE FREEDOM.
P(i) = 1 - PROCCHI(WALD(i),1) P VALUE FOR EACH X(i).

/****** Change this line for TABLES 2 and 3 **/
%MACRO Logistic(x_num, indata, var_y, var_x, drpvar=);
PROC logistic data=&indata outest=temp covout;
model &var_y = &var_x;
RUN;
/****** Change this line for TABLES 2 and 3 **/
Code For TABLE 1

PROC format;
value $pfmt
   'PD' = ' 1 x DIAB'
   'PH' = ' 1 x HTN'
   'PDH' = ' 1 x DIAB and 1 x HTN'
   'PDD' = ' 2 x DIAB'
   'PHH' = ' 2 x HTN'
   'PDDH' = ' 2 x DIAB and 1 x HTN'
   'PHHD' = ' 2 x HTN and 1 x DIAB'
   'PDDHH' = ' 2 x DIAB and 2 x HTN' ;
RUN;

DATA job;
   set man.uc24
     (keep=MMSA_OX PD PH PDH PDD PHH PHHD PDDHH 
      VIAGEO1 SEX D1 D2 D3 D4 WSTHPRO1 BMI01) ;
RUN;
   /*******************************************************************/
   * Call macro LOGISTIC which computes *
   * the odds ratios and CIs *
   * indata = Input Dataset *
   * x_num = # of Independent Variables *
   * y_var = Dependent Variable *
   * x_var = Independent Variables *
   * drpvar = Variables that get Dropped *
   *******************************************************************/
   %Logistic(indata=job, 
      x_num=8, 
      yvar=MMSA_OX, 
      xvar=PD PH PDH PDD PHH PHHD PDDHH VIAGEO1 SEX D1 D2 D3 D4, 
      drpvar=VIAGEO1 SEX DI D2 D3 D4);
   /*******************************************************************/
   * Produce Table 1 *
   *******************************************************************/
PROC tabulate data=good noseps order=formatted;
   CLASS variable;
   VAR or 195 or u95_or ;
   TABLE variable = J 1, 
      (or 195 or u95_or)*5um=l l*f=ltJ.3/ 
      BOX='Nufiber and Type of Disorders*' rts=25;
   LABEL or = 'Odds Ratio' 
      195_or = '95% CI Lower Bound' 
      u95_or = '95% CI Upper Bound';
   FORMAT variable $pfmt. ;
   TITLE2 'TABLE 1';
   FOOTNOTE '* Diabetes (DIAB) and Hypertension (HTN)';
RUN;

Code For TABLE 2

Now that we know we can replicate the results from PROC LOGISTIC, we want to increase the information shown in the table. To do that, we need to add a few lines of code to %LOGISTIC and modify a few other lines.

The macro variable MD (model number) needs to be added to the list of variables.

%LOGISTIC(md=, indata=, x_num=, yvar=, xvar=, drpvar=);

Then we need to specify the length of MODNUM after the "DATA good...." line.

length modnum $6 ;

Assign the value of &MD to MODNUM. This is inserted right before the "output" statement in the "DATA good" section.

modnum = "&md" ;
Need to append all models into the dataset "all", so you have to insert these lines right after the "DATA good" section.

```bash
/**********************
 * Append All Models *
 **********************/
PROC APPEND base=all DATA=test;
run; quit;
```

Those are all the changes and modifications to %LOGISTIC that you need to create TABLE 2. Now let's look at how the code that calls %LOGISTIC needs to be tweaked to produce the desired table.

Since you are appending datasets, it is a good idea to make sure that the dataset "all" does not already exist. Run the following procedure to take care of that.

```bash
PROC datasets lib=work;
delete all;
RUN; QUIT;
```

Then you need to add the variable md to your macro call.

```bash
/****** Additional Variable Needed ******/
* md = Model Number *
*******/
%logistic(md=1**, indata=job,x_num=8,yvar=mmsa_Ox,
  xvar=PD PH PDH POD PHH PDDH PDDH PDDH V1AGE01 SEX D1 D2 D3 D4,
  drpvar=V1AGE01 SEX D1 D2 D3 D4);
%logistic(md=2**, indata=job,x_num=8,yvar=mmsa_Ox,
  xvar=PD PH PDH POD PHH PDDH PDDH PDDH V1AGE01 SEX D1 D2 D3 D4 BM101,
  drpvar=V1AGE01 SEX D1 D2 D3 D4 BM101);
%logistic(md=3**, indata=job,x_num=8,yvar=mmsa_Ox,
  xvar=PD PH PDH POD PHH PDDH PDDH PDDH V1AGE01 SEX D1 D2 D3 D4 BM101,
  drpvar=V1AGE01 SEX D1 D2 D3 D4 BM101);
%logistic(md=4**, indata=job,x_num=8,yvar=mmsa_Ox,
  xvar=PD PH PDH POD PHH PDDH PDDH PDDH V1AGE01 SEX D1 D2 D3 D4 BM101,
  drpvar=V1AGE01 SEX D1 D2 D3 D4 BM101);
```

Finally, update your PROC TABULATE code (highlighted).

```bash
TABLE 2
PROC TABULATE DATA=all noseps ORDER=FORMATTED;
CLASS variable modnum;
VAR or 195 or u95_or ;
TABLE variable = J ,
  modnum='Model Number'*(
    (or 195 or u95_or)*sum=J ~*f=10.3
  ) / rts=25 BOX='Number and Type of Disorders*';
LABEL or = 'Odds Ratio'
  195_or = '95% CI Lower Bound'
  u95_or = '95% CI Upper Bound' ;
FORMAT variable $pfmt. ;
TITLE2 ' TABLE 2';
FOOTNOTE1 '** Diabetes (DIAB) and Hypertension (HTN)';
FOOTNOTE2 *** Adjusted for Age, Gender, Ethnicity/Center';
FOOTNOTE3 **** Adjusted for Age, Gender, Ethnicity/Center, BMI';
FOOTNOTE4 ***** Adjusted for Age, Gender, Ethnicity/Center, WHR';
FOOTNOTE5 ****** Adjusted for Age, Gender, Ethnicity/Center, BMI, WHR';
RUN;
```

And that is all that is needed to create TABLE 2.
Code For TABLE 3

Finally, let's look at what is needed to create a table with two variables going down the left side and two models across. The only changes to %LOGISTIC are adding `mm $8` to the length statement, and inserting the following line:

```
mm = "$yvsr";
```

right before the "output" statement, with both of these going into the "DATA good" section. The only changes/additions needed in the code that calls %LOGISTIC are adding necessary formats,

```
value $mmfmt
'"MMSA_AOD" = 'MM vs All Other Diabetics';
value $phfmt
'P1' = '1';
'P2' = '2';
'P3' = '3 - 4';
```

changing the variables in the macro call statement to reflect what you want in the table,

```
%logistic(md=1*,indata=job,x_num=3,yvar=MMSA_AOD,
xvar=P1 P2 P3 VIAGE01 SEX D1 D2 D3 D4,
   drpvar=VIAGE01 SEX D1 D2 D3 D4);
%logistic(md=1*,indata=job,x_num=3,yvar=MMSA_AOH,
xvar=P1 P2 P3 VIAGE01 SEX D1 D2 D3 D4,
   drpvar=VIAGE01 SEX D1 D2 D3 D4);
%logistic(md=2**,indata=job,x_num=3,yvar=MMSA_AOD,
xvar=P1 P2 P3 VIAGE01 SEX D1 D2 D3 D4 BM101 WSTHPRO1,
   drpvar=VIAGE01 SEX D1 D2 D3 D4 BM101 WSTHPRO1);
%logistic(md=2**,indata=job,x_num=3,yvar=MMSA_AOH,
xvar=P1 P2 P3 VIAGE01 SEX D1 D2 D3 D4 BM101 WSTHPRO1,
   drpvar=VIAGE01 SEX D1 D2 D3 D4 BM101 WSTHPRO1);
```

and finally modify the PROC TABULATE.

```
CLASS mm variable modnum;
TABLE mm="*",
    variable="*",
    modnum="Model Number"*(or 95_or w95_or)*sum="*f=10.3/"
   rle=46 box="Case-Control Contrast Number Of By Type Of Component Histories";
FORMAT mm $mmfmt. variable $phfmt. ;
TITLE5 ' TABLE 3';
FOOTNOTE1 '* Adjusted for Age, Gender, Ethnicity/Center';
FOOTNOTE2 '** Adjusted for Age, Gender, Ethnicity/Center, BMI, WHR';
```

This creates TABLE 3.

Conclusion

So now you know one way to get odds ratios and confidence intervals into PROC TABULATE output. There are other ways to do the same thing, this approach is just one that is easily modified to create the different table looks. In addition, macro variables can be added to take advantage of the different features of PROC LOGISTIC, but that is for another day.

Author Contact

Stephen Noga (uccsmm.cscc@mhs.unc.edu)
Dingyi Zhao (uccdyz.cscc@mhs.unc.edu)
University Of North Carolina - CH Collaborative Studies Coordinating Center
137 E. Franklin Street
Suite #203
Chapel Hill, NC 27514
(919) 962-6971
### TABLE 1

<table>
<thead>
<tr>
<th>Number and Type of Disorders*</th>
<th>Odds Ratio</th>
<th>95% CI Lower Bound</th>
<th>95% CI Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 x DIAB</td>
<td>1.536</td>
<td>1.013</td>
<td>2.329</td>
</tr>
<tr>
<td>1 x HTN</td>
<td>1.422</td>
<td>1.058</td>
<td>1.911</td>
</tr>
<tr>
<td>1 x DIAB and 1 x HTN</td>
<td>2.481</td>
<td>1.743</td>
<td>3.530</td>
</tr>
<tr>
<td>2 x DIAB</td>
<td>3.831</td>
<td>1.304</td>
<td>11.258</td>
</tr>
<tr>
<td>2 x HTN</td>
<td>2.280</td>
<td>1.484</td>
<td>3.501</td>
</tr>
<tr>
<td>2 x DIAB and 1 x HTN</td>
<td>6.357</td>
<td>2.813</td>
<td>14.563</td>
</tr>
<tr>
<td>2 x HTN and 1 x DIAB</td>
<td>4.446</td>
<td>2.820</td>
<td>7.010</td>
</tr>
<tr>
<td>2 x DIAB and 2 x HTN</td>
<td>7.134</td>
<td>2.661</td>
<td>19.123</td>
</tr>
</tbody>
</table>

* Diabetes (DIAB) and Hypertension (HTN)

### TABLE 2

<table>
<thead>
<tr>
<th>Number and Type of Disorders*</th>
<th>Model Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1**</td>
</tr>
<tr>
<td></td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>1 x DIAB</td>
<td>1.536</td>
</tr>
<tr>
<td>1 x HTN</td>
<td>1.422</td>
</tr>
<tr>
<td>1 x DIAB and 1 x HTN</td>
<td>2.481</td>
</tr>
<tr>
<td>2 x HTN</td>
<td>2.280</td>
</tr>
<tr>
<td>2 x HTN and 1 x DIAB</td>
<td>4.446</td>
</tr>
</tbody>
</table>

* Diabetes (DIAB) and Hypertension (HTN)
** Adjusted for Age, Gender, Ethnicity/Center
*** Adjusted for Age, Gender, Ethnicity/Center, BMI
**** Adjusted for Age, Gender, Ethnicity/Center, WHR
***** Adjusted for Age, Gender, Ethnicity/Center, BMI, WHR

### TABLE 3

<table>
<thead>
<tr>
<th>Case-Control Contrast</th>
<th>Number Of By Type Of Component Histories</th>
<th>Model Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1*</td>
</tr>
<tr>
<td></td>
<td>Odds Ratio</td>
<td>95% CI Lower Bound</td>
</tr>
<tr>
<td>MM vs All Other</td>
<td>1</td>
<td>0.909</td>
</tr>
<tr>
<td>Diabetics</td>
<td>2</td>
<td>1.097</td>
</tr>
<tr>
<td></td>
<td>3 - 4</td>
<td>1.725</td>
</tr>
<tr>
<td>MM vs All Other</td>
<td>1</td>
<td>0.980</td>
</tr>
<tr>
<td>Hypertensives</td>
<td>2</td>
<td>1.380</td>
</tr>
<tr>
<td></td>
<td>3 - 4</td>
<td>2.775</td>
</tr>
</tbody>
</table>

* Adjusted for Age, Gender, Ethnicity/Center
** Adjusted for Age, Gender, Ethnicity/Center, BMI, WHR