The Association of Morbid Obesity with Mortality and Coronary Revascularization among Patients with Acute Myocardial Infarction using ARRAYS, PROC FREQ and PROC LOGISTIC

Ashwini S Erande MPH, Shaista Malik MD
University of California Irvine, Orange, California

ABSTRACT
The aim of this study was to investigate the impact of morbid obesity (body Mass index >= 40 kg/m2) on in-hospital mortality and coronary revascularization outcomes in patients presenting with acute myocardial infarctions (AMI). The Nationwide Inpatient Sample of the Healthcare Cost and Utilization Project was used, and 413,673 Patients hospitalized with AMIs in 2009 were reviewed. Morbidly obese patients constituted 3.7% of all patients with AMIs. All the analyses were performed in SAS® 9.3.

ARRAY Statements were used to create the morbid obesity variable based on the ICD9 codes from 24 “Diagnosis” data elements. The SAS procedures PROC FREQ and PROC LOGISTIC were used to perform bivariate and multivariate analyses.

The unadjusted and adjusted analyses performed using PROC FREQ and PROC LOGISTIC respectively, revealed that morbidly obese patients compared with those not morbidly obese were more likely to undergo any invasive coronary procedures when presenting with either ST-segment elevation myocardial infarction and also have a higher mortality rate.

The SAS procedures used to analyze and summarize the data within this context are presented in this paper.

INTRODUCTION
This study has explored the association of worse short-term outcomes for morbidly obese patients presenting with acute coronary syndromes in Nationwide Inpatient Sample (NIS) database of the Healthcare Cost and Utilization Project. The study analyzed the associations among morbid obesity, treatment utilization, and mortality while adjusting for baseline characteristics, including co-morbidities, for 413,673 patients hospitalized with acute myocardial infarctions. This study involved a population-based sample of all patients admitted with AMIs to 1,045 hospitals in 44 states in 2009 whose admission and discharge data were included in the NIS.

We used SAS version 9.3 (SAS Institute Inc., Cary, North Carolina) for all analyses. ARRAY statements were used to create the main diagnosis variable. PROC FREQ and PROC LOGISTIC were used to perform bivariate and multivariate analyses to obtain the chi squares, p-values and Odds ratio with Confidence Intervals respectively.

PROC FREQ with the chi square option was used to find differences between morbidly obese and normal (not morbidly obese) patients with respect to the categorical variables; coronary revascularization procedure utilization and mortality experienced post those procedures and to obtain distributional statistics.

PROC LOGISTIC was performed as unconditional logistic regression to estimate adjusted odds ratios (ORs) for in-hospital mortality as well as procedure use that was performed. To control for differential characteristics of morbidly obese patients and those not morbidly obese, covariates including age, gender, race, income, Elixhauser co-morbidities, and hospital characteristics such as hospital location, hospital control (profit or nonprofit), hospital teaching status, and hospital volume were included in the models. All analyses were weighted using NIS-provided weights to create national estimates.

This paper discusses the above listed SAS procedures and walks through them as they help generate the results that demonstrate how the patients with morbid obesity had higher coronary revascularization procedure use and lower odds of in-hospital mortality, compared to those not morbidly obese, consistent with the phenomenon of the “obesity paradox.”

STUDY SAMPLES AND DATA ELEMENTS
The key interest variable in the study was ‘morbid obesity’. Other explanatory variables in the risk adjustment algorithm include the socio demographic variables, hospital factors and 30 chronic co-morbidities. These are specified in the multivariate analyses section.
PROC FREQ, PROC LOGISTIC and Array statements help a great deal in assessing the effect of Morbid Obesity on In-hospital, continued

The primary sample contained - AMI (Acute Myocardial Infarction) cases. Sub-Samples consisted of STEMI (ST Elevated MI) and NSTEMI (Non-ST Elevated MI) cases

Response Variables: The principal outcome measure was short-term all-cause mortality (in-hospital mortality), which was defined as death that occurred during the initial hospitalization, between the day of hospital admission and date before discharge. Secondary outcomes included utilization of coronary procedure and those were diagnostic coronary angiography, percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) surgery.

These data include ICD-9-CM-coded primary and secondary diagnoses; primary and secondary procedures; admission and discharge status; demographic information such as gender, age, race and ethnicity, and median income for ZIP code divided into quartiles; expected payment source; total charges; length of stay; and hospital region, teaching status, ownership type, and bed size. We used ICD-9-CM secondary diagnosis codes and a database-defined variable for morbid obesity (body mass index ≥40 kg/m²) developed by the Agency for Healthcare Research and Quality. ICD-9-CM secondary codes were used to indicate the presence of up to 30 chronic co-morbidities likely to have been present on admission, using the Elixhauser comorbidity adjustment method developed at the Agency for Healthcare Research and Quality.

PROCEDURES FOR DATA TRANSFORMATION AND BIVARIATE ANALYSES.

In this section I have briefly introduced the Array and Do Loop statements to show their use in creating the main variable of interest that is Morbid obesity.

ARRAY AND DO LOOP STATEMENTS TO CREATE THE 'MORBIDOBESITY' VARIABLE:

Below is the SAS code that was used to include ARRAY Statement and Do Loop to create the Morbid Obesity variable, using 25 data elements that indicated primary, secondary and tertiary diagnoses.

```
data Obesity;
set Obesity;
array dmxmorbidobe{24} dx2-dx25;  
morbidobesity=0;
do i=1 to 24;  
if dmxmorbidobe{i} in ('27801') then morbidobesity=1;
end;
run;
```

The Array is defined as 'dmxmorbidobe' and is assigned 24 members; dx2 to dx25 in step 1 (which are the 'Diagnoses' variables in the dataset) to create the morbidobesity’ variable. Dx1 is not used here as it is a primary diagnosis and is already used to define the main study sample variable called” Acute Myocardial Infarction.

In Step 2, the 'if' statement indicates the condition when the variable 'morbidobesity’ is coded 0 or 1 in step. And the Do loop (do i=1 to 24;) is employed to check if the condition is satisfied for each of the 24 diagnoses variables for every record before assigning 0 or 1 to the newly created morbidobesity’ variable. Thus the newly created variable 'morbidobesity' reads as 1=morbid obesity present and 0=Absence of morbid obesity.

PROC FREQ FOR BIVARIATE ANALYSES:

In bivariate analyses, PROC FREQ: shows the association between 2 categorical variables by providing numbers and percentages.

PROC FREQ:  To compare co-morbidities in between morbidly obese patients and the non-morbidly obese patients.

```
Proc freq data= Obesity;
tables morbidobesity* (race1 income diabetes hypertension renalfail perivasc hyperlipidemia) / chisq;  
weight discwt;  
title 'Morbidobesity by race income and by main comorbidities';
run;
```
In Step ①, the ‘morbid obesity’ variable compared with each of the variables listed in the parentheses separately. The use of option ‘chisq’ generates the chi square and p-value (highlighted in the output 1) that can be reported as baseline results.

In Step ②, the Weight statement uses the variable ‘discwt’ to provide weight to each observation in the input dataset to accomplish the weighted analyses.

Below in Output 1, each cell contains 4 numbers; they are cell frequencies, percentages of the total frequencies, percentages of row frequencies and percentages of column frequencies, from top to bottom respectively. The Row Pct gives the percent of observations in the row; for instance; 5577.19*100 /15254.3 = 36.56%. Col Pct gives the percent of observations in the column; for instance; 5577.19*100 /272757 = 2.04%.

---

The FREQ Procedure
Table of Morbidobesity by Diabetes

<table>
<thead>
<tr>
<th>morbidity</th>
<th>Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td>NO</td>
<td>267180</td>
</tr>
<tr>
<td></td>
<td>97.96</td>
</tr>
<tr>
<td>Yes</td>
<td>5577.19</td>
</tr>
<tr>
<td></td>
<td>2.04</td>
</tr>
<tr>
<td>Total</td>
<td>272757</td>
</tr>
</tbody>
</table>

Statistics for Table of Morbidobesity by Diabetes

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td></td>
<td>6084.3649</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Likelihood Ratio Chi-Square</td>
<td>1</td>
<td>5682.7885</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Continuity Adj. Chi-Square</td>
<td>1</td>
<td>6083.0071</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mantel-Haenszel Chi-Square</td>
<td>1</td>
<td>6084.3502</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Phi Coefficient</td>
<td></td>
<td>0.1213</td>
<td></td>
</tr>
<tr>
<td>Contingency Coefficient</td>
<td></td>
<td>0.1204</td>
<td></td>
</tr>
<tr>
<td>Cramer’s V</td>
<td></td>
<td>0.1213</td>
<td></td>
</tr>
</tbody>
</table>

Sample Size = 413673.09388

Output 1. Portion Of The Output From PROC FREQ With Chi Square

Chi-square - Also known as Pearson chi-square and is of interest in this particular analyses. It compares the difference between the observed frequencies with the expected frequencies collectively (considering the degree of freedom for each of the variables). If the p-value is small enough (say < 0.05), then we will reject the null hypothesis that the two variables are independent and conclude that there is indeed an association between the two variables.

Likelihood Ratio Chi-Square - This compares the ratio between the observed and the expected frequencies and is second most widely used after the chi-square test. It is directly related to log-linear analysis and logistic regression.

Mantel-Haenszel Chi-Square - It is an ordinal measure of significance. Not applicable in this case.
Within the scope of this paper, we will consider only the chi-square and its p value; which are 6084.3649 and <.0001 respectively. These values tell us that there is an association between these two variables and that significantly higher number of cases with morbid obesity have diabetes as compared with those without morbid obesity.

Following two examples are the codes and output of morbid obesity tested against two outcome variables of interest.

1. **PROC FREQ**: To compare any coronary revascularization procedure utilization between morbidly obese patients and the non-morbidly obese patients.

```nc
proc freq data= Obesity;
tables morbidobesity*anyprocedure/chisq;
where stemi=1;
weight discwt;
title 'Chisquare of morbidobesity vs PCI use in STEMI patients';
run;
```

In step 6, the ‘where’ statement is used to define the sub-set of the main input dataset; in this case; using ONLY the STEMI cases to find the association.

The FREQ Procedure
Table of morbidobesity by Anyprocedure

<table>
<thead>
<tr>
<th>morbidobesity</th>
<th>Anyprocedure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td>8041.61</td>
</tr>
<tr>
<td></td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>6.17</td>
</tr>
<tr>
<td>Yes</td>
<td>94.5066</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>2.61</td>
</tr>
<tr>
<td></td>
<td>1.16</td>
</tr>
<tr>
<td>Total</td>
<td>8136.12</td>
</tr>
<tr>
<td></td>
<td>6.07</td>
</tr>
</tbody>
</table>

Statistics for Table of Morbidobesity by Anyprocedure

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>1</td>
<td>78.2662</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Likelihood Ratio Chi-Square</td>
<td>1</td>
<td>97.8717</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Continuity Adj. Chi-Square</td>
<td>1</td>
<td>77.6435</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mantel-Haenszel Chi-Square</td>
<td>1</td>
<td>78.2657</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Phi Coefficient</td>
<td></td>
<td>0.0242</td>
<td></td>
</tr>
<tr>
<td>Contingency Coefficient</td>
<td></td>
<td>0.0242</td>
<td></td>
</tr>
<tr>
<td>Cramer's V</td>
<td></td>
<td>0.0242</td>
<td></td>
</tr>
</tbody>
</table>

Sample Size = 134032.31858

Output 2. The Difference Between Morbidly Obese And Non-Morbidly Obese Cases Presenting With STEMI In Getting Any Coronary revascularization Procedure
The SAS code above produces the result with chi square and p-value to show the difference in use of coronary catheterization between the morbidly obese and not-morbidly obese cases presenting with STEMI.

The Output 2 shows us the unadjusted results and can be interpreted as follows: a significantly higher number of morbidly obese patients were undergoing any invasive coronary procedures when presenting with STEMI (97.4% vs 93.8%, p < 0.0001) compared with those not morbidly obese.

2. PROC FREQ: to compare in-hospital mortality in morbidly obese patients and the non-morbidly obese patients.

The code below produces the results with chi square and p-value to show the difference in mortality in morbidly obese and not-morbidly obese cases presenting with STEMI.

```sas
proc freq data= Obesity;
tables morbidobesity*Died/chisq;
where stemi=1;
weight discwt;
title 'Chisquare of morbidobesity vs Mortality use in STEMI patients';
run;
```

Output 3. Difference In In-Hospital Mortality Between Morbidly Obese And Non-Morbidly Obese Cases.
The Output 3 shows lower unadjusted in-hospital mortality rate in morbidly obese patients compared with those not morbidly obese in the STEMI subsample (4.7% vs 6.3%, p <0.0001).

After running PROC FREQ for rest of the invasive coronary procedures against the 'morbid obesity', separately, we get the same output as shown in Output 2. Similarly, the PROC FREQ for mortality post each coronary procedure against the 'morbid obesity' separately yields the results as shown in Output 3.

These unadjusted results show significantly higher procedure utilization in morbidly obese patients compared to non-morbidly cases in section 1, while the section 2 shows significantly lower in-hospital mortality post coronary revascularization procedures in morbidly obese cases as compared to their counterparts.

### Output 4. Summary tables reporting Chi square and P values, generated by PROC FREQ

<table>
<thead>
<tr>
<th>Coronary Revasc. Procedure</th>
<th>Morbid obesity (AMI patients)</th>
<th>Morbid obesity (STEMI patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Yes</td>
</tr>
<tr>
<td>Any procedure</td>
<td>351,944</td>
<td>88.3</td>
</tr>
<tr>
<td>Diag. cath</td>
<td>88,250</td>
<td>24.7</td>
</tr>
<tr>
<td>PCI</td>
<td>217,492</td>
<td>45.1</td>
</tr>
<tr>
<td>CABG</td>
<td>46,201</td>
<td>18.6</td>
</tr>
</tbody>
</table>

### PROCEDURES- MULTIVARIATE ANALYSES

This section focuses on the multivariate analyses and discusses the PROC LOGISTIC used to perform the adjusted analyses using Logistic Regression and obtain the point estimates along with the p-values.

Logistic regression describes the relationship between a categorical response variable and a set of predictor variables. A categorical response variable can be a binary variable, an ordinal variable or a nominal variable; in this case it's a binary variable. Each type of categorical variables requires different techniques to model its relationship with the predictor variables. For a binary response variable, such as a response to a yes-no question, logistic regression model is a commonly used model.
The next two examples show the independent effect of morbid obesity on any one coronary revascularization procedure utilization and overall in-hospital mortality in cases presenting with STEMI.

**PROC LOGISTIC**

The independent variables adjusted in models include age, gender, peripheral vascular disease, paralysis, other neurologic disorders, chronic pulmonary disease, diabetes mellitus, diabetes mellitus with chronic complications, hypothyroidism, renal failure, liver disease, peptic ulcer disease, acquired immune deficiency syndrome, lymphoma, metastatic cancer, solid tumor without metastasis, rheumatoid arthritis, coagulopathy, weight loss, fluid and electrolyte disorders, chronic blood-loss anemia, iron deficiency anemia, alcohol abuse, drug abuse, psychoses, and hypertension.

**Effect of morbid obesity on coronary revascularization procedure utilization**

The following code and SAS output provide us the adjusted analyses results to show the effect of presence of morbid obesity on getting any coronary procedure.

```sas
proc logistic data = Obesity descending;
  class female (ref= first) morbidobesity (ref=first) diabetes htn_c aids alcohol ANEMDEF arth racel(ref=first) income(ref=first) hosp_location h_contrl(ref=first) hosp_teach bldloss chf chrnlung coag depress drug hypothy liver lymph lytes mets neuro morbidobesity para perivasc psych pulmcirc renfail tumor ulcer valve wghtloss cararrhythmia/param=ref;
  model anyproc1= age female diabetes htn_c aids alcohol ANEMDEF arth racel income hosp_location h_contrl hosp_teach TOTAL_DISC bldloss chf chrnlung coag depress drug hypothy liver lymph lytes mets neuro morbidobesity para perivasc psych pulmcirc renfail tumor ulcer valve wghtloss cararrhythmia;
  where stemi=1;
  weight discwt;
  title 'Logi Reg anyproc vs Morbid obesity in STEMI cases';
  run;
  quit;
```

This code produces the point estimates and p values that are shown in Output 5.

**Descending** option is specified in the PROC LOGISTIC statement, to indicate which event to model. In this case the ‘anyproc1= 1/yes’ is modeled as an event. The other way of specifying the event to model is to use the 'event' option in MODEL statement using the quotes in the option as event = '1'.

The note is shown in Output5, about which event is modeled as **Probability modeled is “anyproc1” = 1**.

In SAS, the **CLASS** statement is useful in creating the dummy variables for a categorical variable on-the-fly. There are various coding schemes from which to choose. The default coding for all the categorical variables in proc logistic is the effect coding. Here we use dummy coding by using the “**param = ref**” option and by specifying the comparison group by using ref = option after the variable name. For instance, for ‘morbid obesity' variable, the reference category used is “first” which is coded as “0” indicating those cases with absence of morbid obesity.

In Output 5, the first box shows the number of Observations Read and Number of Observations Used - The Number of Observations Used may be less than the Number of Observations Read if there are missing values for any variables in the equation. By default, SAS does a listwise deletion of incomplete cases.

The second box shows the response Variable profile and number of response Levels.

**Parameter** – They are the predictor/explanatory variables in the model and the intercept.

**Estimate** The estimates help here to understand the statistical aspect of the regression equation. The logistic regression models the log odds of a positive response (probability modeled is anyproc1=1) as a linear combination of the predictor variables. This is written as

\[
\log\left(\frac{p}{1-p}\right) = b_0 + b_1 \text{female} + b_2 \text{morbidobesity} + b_3 \text{DM},
\]

where \( p \) is the probability that Anyproc1 is 1. For our model, we have,

\[
\log\left(\frac{p}{1-p}\right) = -8.16 + 0.3107 \text{female} + 0.8112 \text{morbidobesity} + 0.2384 \text{DM}.
\]
PROC FREQ, PROC LOGISTIC and Array statements help a great deal in assessing the effect of Morbid Obesity on In-hospital, continued

We can interpret the parameter estimates as follows: for a one unit change in the predictor variable, the difference in log-odds for a positive outcome is expected to change by the respective coefficient, given the other variables in the model are held constant.

**MorbidObesity** - This is the estimate logistic regression coefficient comparing cases with morbid obesity (‘1’) with those without morbid obesity (‘0’). For a one unit change in MorbidObesity (from moving from 0 to 1), given the other variables in the model are held constant, the difference in log-odds for “anyproc” is expected to increase by 0.8112 unit, given the other variables in the model are held constant.

**Point estimate** - The interpretation for the point estimate of MorbidObesity will be that for a one unit increase (from moving from morbid obesity (1) to No-morbid obesity (0), the odds of getting a “anyproc” increases by (2.25-1)*100% = 125%. For female, from going from male to female (0 to 1), the odds of getting a “anyproc” decreases by (1-0.73)*100% = 27%. The explanation for the female predictor is slightly different than the one for morbidobesity because the point estimate (Odds ratio) is less than 1, and actually renders the female gender in AMI patients as protective to undergo any cauterization procedure.

**95% Wald Confidence Limits** - For a given predictor variable with a level of 95% confidence the CI’s would include the “true” population odds ratio. If the CI includes one, we’d fail to reject the null hypothesis that a particular regression coefficient equals zero and the odds ratio equals one, given the other predictors are in the model.

<table>
<thead>
<tr>
<th>Probability modeled is anyproc1='1'</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analysis of Maximum Likelihood Estimates</strong></td>
</tr>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>AGE</td>
</tr>
<tr>
<td>FEMALE</td>
</tr>
<tr>
<td>DM</td>
</tr>
<tr>
<td>Morbidobesity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Odds Ratio Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
</tr>
<tr>
<td>AGE</td>
</tr>
<tr>
<td>FEMALE 1 vs 0</td>
</tr>
<tr>
<td>DM 0 vs 1</td>
</tr>
<tr>
<td>Morbidobesity 1 vs 0</td>
</tr>
</tbody>
</table>

Output 5. Portion Of The Output From PROC LOGISTIC: Effect of morbid obesity on coronary revascularization procedure utilization

Output 5 displays part of the PROC LOGISTIC output. It can be interpreted as below:
In those presenting with STEMI's, the likelihood of getting any procedure done was significantly higher in morbidly obese patients as compared to their counterparts. (OR 2.25 95% CI 1.816-2.789) The CI does not include one and the p-value of <0.0001 confirms that we need to reject the null hypothesis and proves that 'morbid obesity' does have an independent predictive effect on coronary procedure utilization.

Effect of morbid obesity on in-hospital mortality

This code provides us the adjusted results to show the effect of presence of morbid obesity on in-hospital mortality.

```
proc logistic data =Obesity descending;
class female (ref= first) morbidobesity (ref=first) diabetes htn_c aids alcohol ANEMDEF arth racel(ref=first) income(ref=first) hosp_location h_contrl(ref=first)
hosp_teach bldloss chf chrnlung coag depress drug hypothy liver lymph lytes mets neuro morbidobesity para perivasc psych pulmcirc renlfail tumor ulcer valve wghtloss cararrhythmia/param=ref;
model Died= age female diabetes htn_c aids alcohol ANEMDEF arth racel income hosp_location h_contrl hosp_teach TOTAL_DISC bldloss chf chrnlung coag depress drug hypothy liver lymph lytes mets neuro morbidobesity para perivasc psych pulmcirc renlfail tumor ulcer valve wghtloss cararrhythmia;
weight discwt;
title 'Logi Reg Mortality vs Morbid obesity in AMI cases';
runt;
```

Output 6. Portion Of The Output From PROC LOGISTIC: Effect of morbid obesity on in-hospital mortality

In this PROC LOGISTIC, the event is modeled as Probability modeled is “died” =1/yes.
PROC FREQ, PROC LOGISTIC and Array statements help a great deal in assessing the effect of Morbid Obesity on In-hospital, continued

For the overall sample presenting with AMI, the in-hospital mortality was 13 % lower for morbidly obese patients as compared to those who were not morbidly obese (OR 0.87, 95% CI 0.78-0.94). Similar results were observed in STEMI and NSTEMI cases.

In the Output 7, the summary tables show the compiled results obtained by PROC LOGISTIC. The section 1 in Output 7 shows the odds ratio, confidence intervals and P values, obtained by modeling each invasive coronary procedure in a separate PROC LOGISTIC. The section 2 provides the same parameters obtained by modeling positive response for mortality post each coronary revascularization procedure in morbidly obese patients compared to their counterparts.

Thus the adjusted analyses using PROC LOGISTIC shows that the morbidly obese cases are significantly more likely to undergo invasive coronary procedures (except PCI) compared to the non-morbid cases as shown in section 1. The section 2 actually validates the Obesity paradox by displaying significantly lower likelihood of mortality in morbidly obese cases compared to their counterparts.

Output 7. Summary tables reporting Odds ratio, CI and P values, generated by PROC LOGISTIC
CONCLUSION

In any analyses, knowing the data is the first step. PROC FREQ is a very useful tool that can be used in exploring the basic information, providing the distributional analyses and reporting the descriptive statistics for the categorical variables. Upon using the chi-square option it also generates the p-values that can be reported as a part of the bivariate results.

PROC LOGISTIC is a powerful test that describes the relationship between a categorical response variable and a set of predictor variables and it generates the point estimates with the confidence interval. Both these procedures are easy and effective in obtaining bivariate and multivariate analyses results as well as confirming the pre-existing Obesity paradox.

ACKNOWLEDGMENTS

This paper is based on the paper listed below in the recommended readings and which I was a part of. I would like to thank the authors Dhoot Jasdeep, Tariq Shamail, Alpesh Amin, Pranav Patel and Shaista Malik, for their significant contribution in the original paper.

I would also like to thank all the SAS experts from the SAS discussion forum for their valuable comments and suggestions while developing the SAS code.

- Institute of Digital Research Education UCLA; http://www.ats.ucla.edu/stat/sas/seminars/sas_logistic/logistic1.htm

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Ashwini S Erande  
Enterprise: University Of California Irvine  
Address: 101 The City Drive South  
City, State ZIP: Orange, CA, 92868  
Work Phone: 714-456-7025  
Fax: 714-456-7604  
E-mail: aerande@uci.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.