

Paper 212-2009

Using SAS® to Explore the Impact of a Care Management Maternity Program

Barbara B. Okerson, HMC, Richmond, VA

ABSTRACT

Pregnancy and childbirth-related conditions make up almost 25 percent of hospitalizations in the United States. Mothers-to-be that are healthy and well-informed about their needs and the needs of their baby have healthier babies and lower costs. This paper uses SAS statistical procedures combined with ODS statistical graphics and SAS/GRAPH procedures to produce graphics that can be used to analyze the need for, and effectiveness of, a prenatal care-management program that is designed to reduce cost while reducing incidence of preterm and low birth weight babies, especially for high-risk pregnancies.

In addition to a broad range of statistical procedures, SAS/STAT provides customizable statistical graphics that display automatically when ODS graphics are invoked; these are especially effective for graphically displaying the effects and safety of treatments in a mode that is often more easily understood and remembered than tabular forms. Types of data used in this analysis include medical, laboratory, and pharmacy claims combined with demographics and self-reported health readiness assessments (HRAs). Results included in this paper were created with version 9.1.3 of SAS on a Windows XP platform and use Base SAS, SAS/STAT and SAS/GRAPH. SAS Version 9.1 or later is required for ODS graphics extensions. The techniques represented in this paper are not platform-specific and can be adapted by both beginning and advanced SAS users.

INTRODUCTION

Health Management Corporation (HMC) is one of the nation's largest, most experienced providers of integrated care and total health solutions. Since 1983, HMC has offered members and clients comprehensive programs and services intended to empower members to take control of their health and see positive results. Through a population-based approach, HMC's disease management programs provide solutions for prevention, chronic condition support, lifestyle management and complex condition care. HMC's mission is to help improve the health and financial outcomes of its clients and members through innovative health solutions that consider every single member at his/her level of care.

Within HMC, the informatics team provides a broad range of analytical and data services, including information management, research, advanced analytics, client outcomes, business consulting, and data management that ensure the effectiveness and client satisfaction for the services provided as well as add to general knowledge of the health management arena. SAS is the major analytic software used by HMC and, as such, all of the graphics in this paper were produced with SAS software in conjunction with this work.

Data sets used in this paper are either publicly available or are test data sets developed for the purpose of data modeling and technique demonstration and are not intended to represent the health statistics of an actual population or account.

APPLICATION

Pregnancy and delivery combine as major sources of hospitalization costs in the US commercial segment. The high rate of low birth weight infants is especially costly, both in terms of human life (mortality) and in economic terms. In 2006, the low birth weight rate was 8.3 percent of all births; this rate has risen by 19 percent since 1990. Adding to that cost are preterm births which has overlap with low birth weight infants. The 2006 preterm rate was 12.8 percent, a 21 percent increase since 2000 (Hamilton, 2007).

The HMC care management program for pregnancy is designed to help lower the incidence of preterm and low-birth-weight births, reduce pregnancy-related medical costs, and improve the ability of mothers-to-be to recognize the signs and symptoms of complications. The program includes:

- An assigned prenatal/postnatal nurse who is supported by an on-site team of clinical specialists;
- A care management system that assesses risks to help ensure consistent, individualized education and support;
- Prenatal interventions based on the expectant mother's pregnancy status, complications, medical history and physician's plan of care;
- A 24-hour, toll-free nurse access line that helps expectant parents make informed health care decisions;
- A maternity Web site with easy access to educational resources as well as online enrollment and interaction with HMC maternity nurses;
- Targeted support for smoking cessation;
- Postpartum depression screening.

Three types of graphics are included in this paper. The first set (examples 1-2) showcases methods of visualizing the scope of factors to be addressed by the program. The second set (examples 3-5) presents methods for visualizing the results of the care management program. The first set can be used to demonstrate the need for the program, the second to demonstrate the value. When reformatted to show baseline/remeasurement, the "need" graphics can also be used to show value. The last two (examples 6-7) are diagnostics used to check data and assumptions for analysis. Included graphics look at the following:

- o Low birth weight rates;
- o Delivery type by risk factor;
- o Satisfaction survey histogram;
- o Arrow chart of progress in cost risk reduction;
- o Geographical distribution;
- o Survival analysis residuals for mother's weight vs. birth weight; and
- o Box plot of maternity weight gain.

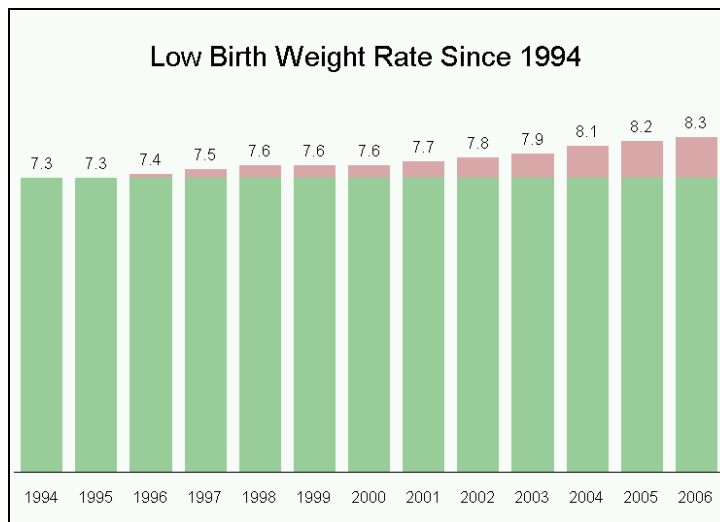
Selected SAS program code is included with each graphical display. Full SAS code for each graphic is available by request from the author.

EXAMPLE 1: STACKED BAR CHART WITH DATA VALUES.

A bar chart is a graphical display of data that includes rectangular bars with lengths proportional to the values that they represent. A stacked bar chart combines subgroups of categories into single bars for each category. Bars can be vertical or horizontal; however, these charts are only useful for discrete data. Edmund Tufte, the graphics guru, has suggested that stacked bar charts are one of the forms of graphics to be avoided (Tufte, 2001). His reasoning is that stacking the values makes it very difficult to estimate the value represented on the top of the bars. The graphic below gets past this objection by including the total value as part of the graphic and changing the color of the portion of the bar that increased from the base value. There is no longer any difficulty in estimating the value of the top of the bar; all of the values can be accurately calculated from the information displayed.

Example 1 Graphic – Bar chart depicting change in rate of low birth weight.

Low birth weight infants represent only a small number of all births, but a major portion of the annual cost of pregnancy and childbirth. The graphic below uses data from the Centers for Disease Control and Prevention to show the increase in low birth weight births from 1994-2006. This presents a challenge for care management programs to help pregnant and pre-pregnant women get the information and care that they need to reduce the number and cost of low birth weights.



Example 1 Selected SAS Code:

The data used in this example, reads in the values for each year; then the difference between the 1994 value and the current value is calculated. A new variable is created to distinguish between the base value and the difference value for each year. Unnecessary labels and lines are suppressed from the graph so that the information (increase) is readily apparent to the viewer.

```

/*Set Titles, Patterns and Axes*/
pattern1 color=CX99CC99;
pattern2 color=CXDF9A99;
axis1 label=none minor=none major=none value=none style=0 order=(0 to 8.5 by .5);
axis2 label=none;
Title 'Low Birth Weight Rate Since 1994';

```

```

/*Create the Stacked Bar Chart*/
proc gchart data=lbw;
  vbar year/discrete sumvar=rate width=10 noframe subgroup=category
  raxis=axis1 maxis=axis2 outside=sum nolegend;
run;
quit;

```

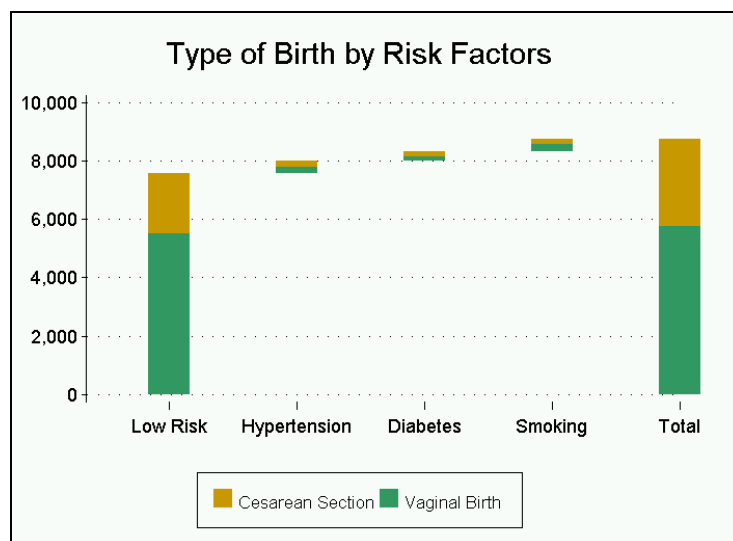
EXAMPLE 2: WATERFALL CHART.

The waterfall chart is ideal for displaying areas of loss or gain between starting and ending values. In general use, waterfall charts show how an initial value is increased and decreased by a series of intermediate values, leading to a final value. The advantage of displaying this type of information in a waterfall chart over a stacked bar chart is that, by isolating the subcategories of interest, the viewer's eye is drawn to the highlighted information. The waterfall chart is a relatively new graphic that has recently been popularized by McKinsey and Company for illustration of quantitative flows.

In this example, the waterfall chart shows a standard risk population, with the intermediate values as high risk populations. This type of graph is often accomplished by creating an invisible column that keeps the increases and decreases linked to the heights of the previous columns. This method is the standard for Excel and is also demonstrated in a recent SESUG paper using Proc GHART (Booth, 2008). This example used Proc GPLOT with annotate and does not require computation of invisible columns, nor does it require the floating columns to be based on values of other columns. SAS Graphics developers have a waterfall chart planned in a future graphics release.

Example 2 Graphic – Waterfall chart of birth type by risk factor.

Part of understanding a high risk maternity population involves isolating differences in cost factors, e.g. cesarean vs. vaginal birth. In the graphic below the proportion of vaginal to cesarean birth is greater for the low risk population than the risk population (more cesarean births in the risk populations).



Example 2 Selected SAS Code:

The following example code includes creation of the Annotate dataset and the GPLOT procedure. The waterfall dataset used by this example includes the following variables: Risk, Type, Begin, End, and Line. The bars are created in the annotate data set, which allows the beginning and ending points to be data driven. Because the bar is wider than a data point, it is necessary to use the offset option in the horizontal axis statement to move the starting bar to the right; otherwise the bar will overlap the vertical axis line. To add additional categories and colors, just add these to the color statement in the annotate data step.

```

/*Create the annotate data set*/
data anno;
  length function color $8;
  retain xsys ysys '2' size 40 color 'cx66CC33' when 'a';
set waterfall;
  line=line;
  if type="C" then color='cxFF9900';
  else color='cx66CC33';
  function='move'; y=begin; xc=Risk; output;
  function='draw'; y=end; xc=Risk; output;
run;

/*Set Titles, Symbols and Axes*/

```

```

title "Type of Birth by Risk Factors";
symbol1 i=none;
axis1 order=("Low Risk" "Hypertension" "Diabetes" "Smoking" "Total")
      offset=(10,0) label = none style=0 ;
axis2 Order=(0 to 10000 by 2000) minor=none label=none;

/*Plot the Timeline*/
proc gplot annotate=anno;
plot end*Risk/haxis=axis1 vaxis=axis2 autovref lvref=34 hminor=0;
format end comma6.;
run;
quit;

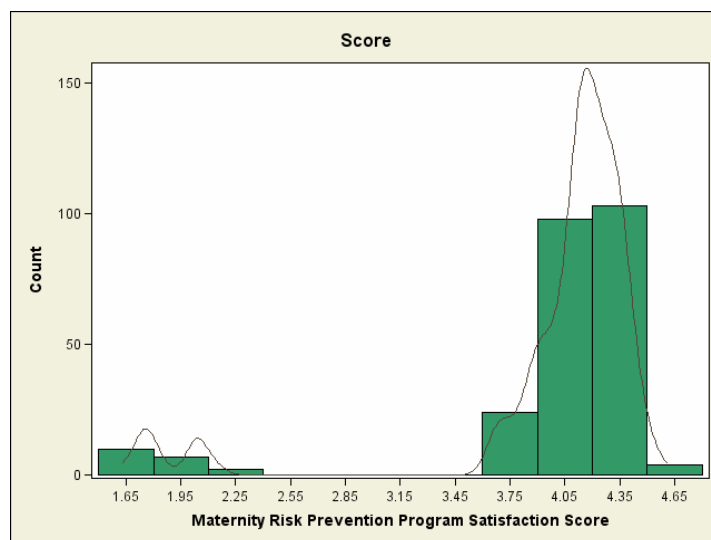
```

EXAMPLE 3: KERNEL DENSITY ESTIMATOR HISTOGRAM.

A histogram is a graphical display of tabulated frequencies, usually shown as a bar graph. In their simplest form, kernel density estimators are smoothed histograms. While a histogram restricts observations to only one group or bin, kernel estimators allow observations to be included in more than one group. It can be thought of as a window that slides over the data. The kernel function weights the data passing through the window on its relative distance to the center of the window. Smoothness is determined by the window interval. Because satisfaction scores are averaged across a number of factors and are therefore continuous, this provides a better view of this distribution than a regular histogram that forces categorical assignment.

Example 3 Graphic – KDE Histogram of Maternity Program Satisfaction Scores.

In the graphic below, the histogram clearly shows that while the vast majority of program participants were satisfied with the program (5='Very Satisfied'), there were a small class of members who may have been unsatisfied. Although there is also the possibility that some participants reversed the scale, comments provided by these survey participants could be used to improve the program.



Example 3 Selected SAS Code:

This output is an example of ODS statistical graphics. Available with selected statistical and ETS procedures, these options allow for the creation of graphics without the necessity of saving data points to an output data set. The graphics use templates and style elements to create presentation-ready graphics. To modify the ODS graphics output, modifications need to be made to the template creating the graphic using Proc TEMPLATE. The template can be located by using the ODS TRACE statement, which returns to the SAS log a listing of output objects created. The graph above is created with the Stat.KDE.Graphics.HistDensity object. Once a template name is located, the template code can be located by typing *odst* (or *odstemplates*) on the command line. From the template window, the specific object code was found in the subdirectories under *sashelp.tmplmst*. The Proc TEMPLATE code as modified for the above graphic is shown below, together with the Proc KDE code.

```

proc template;
define statgraph Stat.KDE.Graphics.HistDensity;
dynamic _HISTOGRAM _DENSITY _DEPLABEL;
layout Gridded;
  layout overlay / padbottom=5;
  layout gridded / columns=3 valign=top;
  entrytitle _DEPLABEL /fontsize=15 padbottom=5 padtop=5;
endlayout;
endlayout;
layout Overlay/XAXISOPTS=(LABEL='Maternity Risk Prevention Program Satisfaction Score') ;

```

```

if (_HISTOGRAM=1)
  HistogramParm x=BINX y=BINCOUNT / fillcolor=cx339966;
endif;
if (_DENSITY=1)
  series x=DENSITYX y=DENSITYCOUNT / linecolor=
  StatGraphFitLine:contrastcolor;
endif;
EndLayout;
EndLayout;
end;
run;

ods noproctitle;
ods html path='r:\bokerson\SASGlobalForum\SGF 2009\' body='kde2009.htm'
  style=statistical;
ods graphics on;

proc kde data=satisfaction;
  univar Score /plots=histdensity bwm=.8;
run;
ods graphics off;
ods html close;

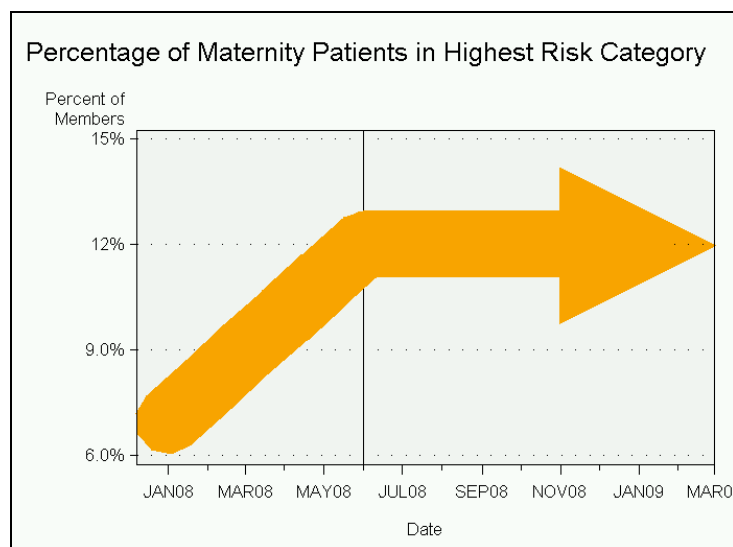
```

EXAMPLE 4: ARROW CHART.

The arrow chart is most often used by print media as a quick depiction of changed direction for a major issue such as the economy, sales and other topics. In reality, the arrow chart is a plot that has been enhanced to show direction. The plot line is widened, a smoothing algorithm is applied to the actual data points, and the points themselves are suppressed.

Example 4 Graphic – Arrow Chart of Intervention Impact with Maternity Patients.

As mentioned previously, high-risk pregnancies are a major source of rising health care cost. In this example, the number of patients in the highest cost risk category had continued a gradual increase. The arrow chart shows that, after intervention, the previous rise in the number of maternity patients in the highest risk category leveled off. The vertical line represents the beginning of intervention.



Example 4 Selected SAS Code:

Because this graphic was created in version 9.1 of SAS, arrows such as this can not be created directly through either graphic procedures or annotate functions. The arrow in this graphic was created using annotate to place a triangle at the end of a thick line represented by the last data point. The horizontal axis is extended to provide space for the arrowhead. The code below creates the annotate data set used in the graphic above. Version 9.2 of SAS will bring more options with the %arrow annotate macro function (run in conjunction with the %annomac macro). The %label macro function takes the following options: x, y, text-string, color, angle, rotate, size, style, and position.

```

data anno;
set plot;
%dc1anno;
%annomac;
%system (2,2);

```

```

if date='1jan09'd then do;
%label (date, percent_of_members, 'B', orange, 0, 0, 10, 'Marker', 5);
end;
run;

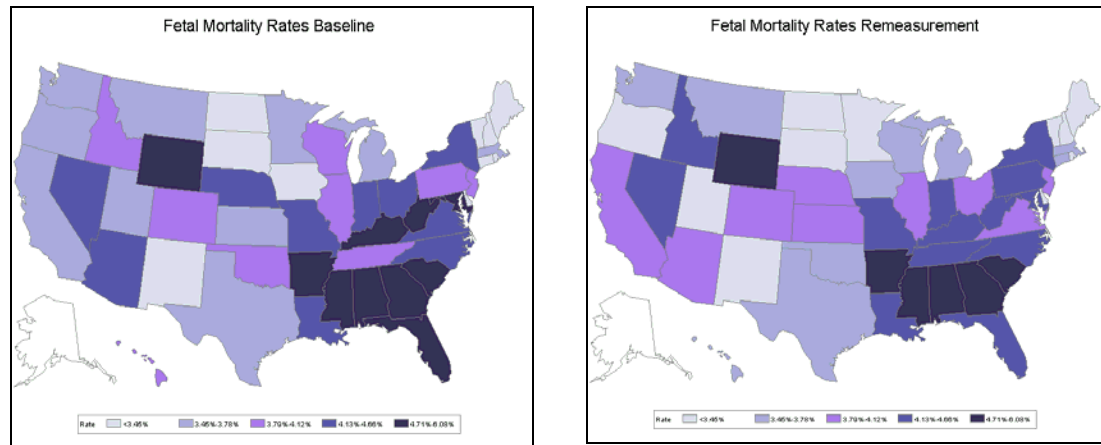
```

EXAMPLE 5: BEFORE/AFTER CHOROPLETH MAP.

A choropleth map is a thematic map in which areas are shaded in proportion to the measurement of the variable being displayed on the map. It provides an easy to see visualization of how the measurement varies across geographic areas. Colors or shading usually form a progression in response to the levels of the mapping variable. A set or series of maps can be used to show change over time as in the example below where changes in fetal mortality rates are depicted. While the data depicted below is from the CDC, similar data could be used at state, county and zip code level to show change. Choropleth maps are produced in SAS/GRAPH with Proc GMAP.

Example 5 Graphic: Choropleth Maps of Fetal Mortality Rates.

One of the purposes of a prenatal health program is lowering the risk of fetal mortality through better education and use of health services. In these maps, states are divided into quintiles for the baseline period. The remeasurement map uses those same quintiles for mapping the later data. Darker colors are higher rates. Note that a number of states improved during the time period depicted.



Example 5 Selected SAS Code:

Proc Rank was used in this example to produce the quintiles for the map. The break points were output and used as the breakpoints for the second map. The code excerpt below shows the goptions (graphic options) statement with the graduated color scheme, the legend statement, and the GMAP procedure code.

```

goptions colors=(white BWH VPAB VLIPB LIB MOPB) ftext='Arial' ctext=black;
legend cborder=gray label=("Rate") value=(j=1 "<3.45%" "3.45%-3.78%" "3.79%-4.12%"
"4.13%-4.66%" "4.71%-6.08%");
title3 h=2 'Fetal Mortality Rates Baseline';
proc gmap map=maps.us data=datamap;
id state;
choro raterank / legend=legend coutline=gray;
run;
quit;

```

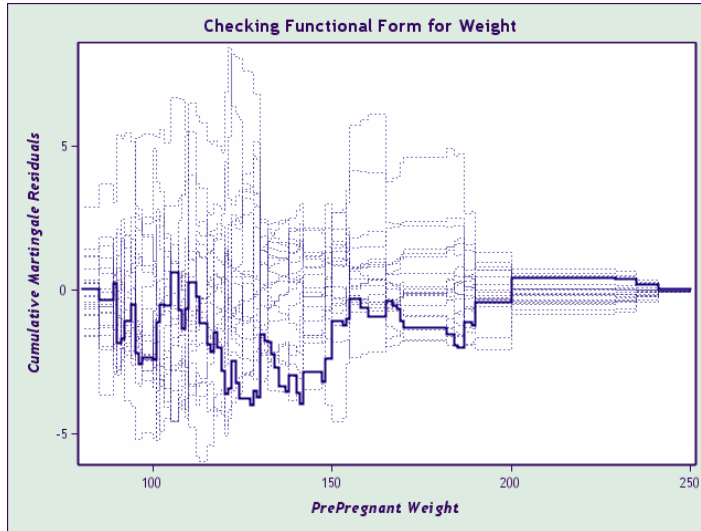
EXAMPLE 6: SURVIVAL ANALYSIS MARTINGALE RESIDUAL PLOT.

Data that measure length of time until the occurrence of an event or measure the lifetime of an event are often referred to as survival data. Survival data variables include a response variable that measures the duration of time until the specified event occurs and can also include independent variables thought to be associated with the failure time variable, e.g. sex, race, age. The event of interest often occurs in a biological setting, as in most medical data, but can also occur in a physical setting, such as engineering data.

The purpose of survival analysis is to model the distribution of the failure time variable and assess the relationship between the failure time variable and the independent variables. SAS provides three procedures for survival analysis; LIFETEST, LIFEREG, and PHREG. Use of relative risk regression models for censored survival data requires collection of covariate information on all individuals under study even when only a small fraction of them die or get diseased, or in the case of this example, give birth to low birth weight infants. For cohort data, methods based on martingale residuals are useful for assessing the fit of a model, as it is important to know the variance of observed versus expected values when determining the validity of the model and data.

Example 6 Graphic – Martingale residuals for mother's weight and low birth weight.

In this example, the PHREG procedure is used as it is most appropriate for data that may be right censored. Censored data is data in which the true value is replaced by some other value because the data does not include the complete range or lifetime. The data set includes several variables that are considered for their influence on low birth weight and includes women between the ages of 16 and 30 who gave birth during the study time period. The plot below shows the observed cumulative Martingale residual process for mother's pre-pregnant weight (solid line) together with simulated realizations from the null distribution. The Martingale residuals represent the difference between the expected number of low birth weights at each point and the observed number of low birth weights.

**Example 6 Selected SAS Code:**

This graphic is another example of a statistical procedure that uses ODS statistical graphics. The graphic itself was created with the default output template. The style (color scheme) was changed on the ODS HTML statement from default to analysis. The graph above uses the Stat.Phreg.Graphics.CResidual object code located in the subdirectories under sashelp.tmplmst.

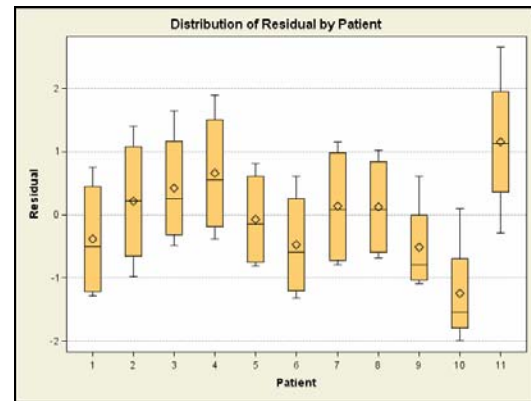
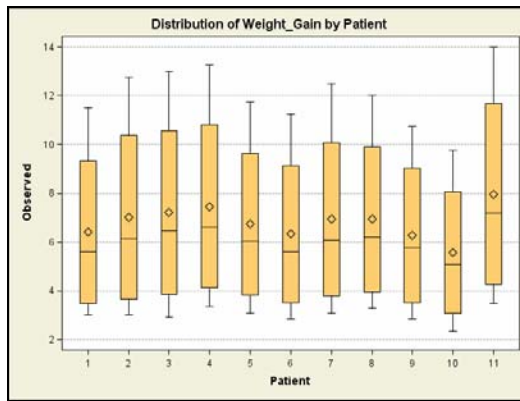
```
ods html style=analysis;
ods graphics on;
proc phreg data=LBW;
  model Time*Low(0)= Weight Comorbids Race Demog/ties=discrete;
  assess var=(weight);
  label weight="PrePregnant Weight";
run;
ods graphics off;
ods html close;
```

EXAMPLE 7: BOX PLOT OF DATA DISTRIBUTION AND RESIDUALS.

A box plot provides a means of graphically depicting groups of numerical data by depicting at a minimum five points: the smallest observation, Q1, median, Q3, and the largest observation. The spacing between different parts of the box plot provide valuable insight into the data, depicting degree of dispersion (spread) and skewness in the data, and identifying outliers. Within Proc MIXED, box plots for effects in the model that comprise a single classification variable are requested with the ODS GRAPHICS statement and the BOXPLOT option of the PROC MIXED statement.

Example 7 Graphic – Box Plot for pregnancy weight gain data distribution.

This example looks at weight gain during pregnancy for 11 patients at several time intervals. The specific box plot options request a plot of the observed data, the marginal residuals, and the conditional residuals. Box plots of the observed values show the variation within and between patients clearly. Because the residuals are only marginally centered on zero, this suggests that time period alone may not be adequate to explain the differences in weight gain.



Example 7 Selected SAS Code:

This output is also an example of default ODS statistical graphics. The nlpnl option is not needed here since there are only 11 observations meeting the criteria. It is provided for example only and can be used to limit the number of box plots in the window.

```
ods html style=statistical;
ods graphics on;
proc mixed data=pgf method=ml boxplot(nlpanel=11);
class Patient;
model Weight_Gain=TimePeriod;
random intercept TimePeriod / type=un subject=Patient;
run;
ods graphics off;
ods html close;
```

CONCLUSION

Visual depictions of data relationships assist analysts and other users in better understanding of data through the quick identification of true associations, anomalies, and areas for further study. In the examples in this paper, SAS software---especially SAS/STAT, SAS/GRAPH and the addition of ODS statistical graphics---provided the flexibility and ease of use allowing for visual identification of data relationships; this information provided a foundation for further analyses to identify additional interactions and critical factors that can enhance the evaluation/effectiveness of a care management program for future moms.

REFERENCES

Alexander GR, Kogan M, Bader D, Carlo W, Allen M, Mor J. "US Birth Weight/Gestational Age-Specific Neonatal Mortality: 1995–1997 Rates for Whites, Hispanics, and Blacks," *Pediatrics*. 2003; 111(1): e61–e66.

Booth, RB. "Ride the Wave – Using Waterfall Graphs to Easily Depict Gaps between Starting and Ending Values." *SESUG Proceedings*. 2008.

Borgan, Ø. and Langholz, B. "Using martingale residuals to assess goodness-of-fit for sampled risk set data." Chapter 4 in *Advances in statistical modeling and inference. Essays in honor of Kjell A Doksum*. Ed.V. Nair. World Scientific, 2007.

Devarajan R. "Need Java Graphics? Use of BIP Graph Components," *SUGI Proceedings*, 2005.

Hamilton BE, Martin JA, and Ventura SJ. "Births: Preliminary Data for 2006". *National Vital Statistics Reports*; vol 56(7). Hyattsville, MD: National Center for Health Statistics. 2007.

Health Management Corporation. <http://www.choosehmc.com>. Accessed January 2009.

McDorman MF, Kirmeyer S, "Fetal and Perinatal Mortality, United States, 2005," *National Vital Statistics Report*, vol 57(8). Hyattsville, MD: National Center for Health Statistics. 2009.

Raisal EM. *The McKinsey Way: Using the Techniques of the World's Top Strategic Consultants to Help You and Your Business*. McGraw-Hill Professional, 1999.

Rodriguez R. "An Introduction to ODS for Statistical Graphics in SAS 9.1," *SUGI Proceedings*, 2003.

SAS Institute Inc. ODS Graphics Overview: ODS Graphics Style Elements and Attributes
<http://support.sas.com/rnd/base/topics/statgraph/proctemplate/a002791342.htm>.
Accessed January 2009.

SAS Institute Inc. 2008. SAS OnlineDoc® 9.1.3. Cary, NC: SAS Institute Inc.

SAS Institute Inc. Output Delivery System. <http://support.sas.com/rnd/app/da/new/daods.html>. Accessed January 2009.

Tufte ER. *The Visual Display of Quantitative Information: Second Edition*. Cheshire, CT: Graphics Press, 2001.

ACKNOWLEDGEMENTS

I would like to acknowledge the members of the HMC outcomes teams and the HMC analytic and research group for their suggestions and assistance in the development of this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. For more information contact:

Barbara B. Okerson, Ph.D., CPHQ
Senior Health Information Consultant,
National Accounts Outcomes
Health Management Corporation (HMC)
8831 Park Central Drive, Suite 100
Richmond, VA 23227
Office: 804-662-5287
Fax: 804-662-5364
Email: bokerson@choosehmc.com



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.