Focusing healthcare quality improvement dollars: Using SAS® for geographic targeting

Barbara B. Okerson, Ph.D., Charlotte F. Carroll, M.S., Virginia Health Quality Center, Glen Allen, VA

ABSTRACT

Like many organizations, the health care quality improvement industry is charged with doing more with less. This can only be accomplished by careful scheduling of resources. Geographical differences in adherence to clinical guidelines have been demonstrated. In this paper, geographical analyses are used both to identify regions and providers for healthcare quality improvement targeting and to identify best practices in both the inpatient and outpatient settings. SAS, Output Delivery System (ODS) and SAS/GRAPH are used to generate the output. All output is created using SAS software, version 8.2 on a Windows 98 platform but are not intended to be platform specific. All SAS code is included in the paper.

INTRODUCTION

The Virginia Health Quality Center (VHQC), winner of the2002 U.S. Senate Productivity and Quality Award, is a health care quality improvement organization that assists healthcare providers in making successful and meaningful changes in the way care is delivered and in improving outcomes of that care, especially for the Medicare community of Virginia. Services provided by VHQC include assisting healthcare organizations with performance improvement techniques, reviewing health plan denials, statistical consulting and data analysis services, health education, credentials verification, clinical and claim database engineering, health care utilization and quality reviews, and coding/DRG validations. VHQC currently collaborates with all acute care and critical access hospitals and invites all physician offices, home health agencies and nursing homes in Virginia to participate in the improvement of health care quality.

In an effort to maximize available resources, geographic analysis is used both to (1) identify regions, and providers within regions, most in need of improvement for defined clinical quality indicators and (2) identify locations of excellence from which opinion leaders can be recruited. From within defined clinical topics, areas for quality improvement focus are selected for regions based on overall lower rates relative to other regions. Speakers are recruited within each region based on high rates and successful implementation of a quality improvement project. This methodology is used both in inpatient and outpatient settings.

The SAS software system provides a variety of tools for spatial data analysis and display. Geographic data display methods include the integrated use of maps, box plots, pie charts, and tables, as well as a multi-graph display technique. This paper focuses on the methods for these analyses.

INPATIENT

Clinical areas have been identified by the Centers for Medicare & Medicaid Services (CMS) for quality improvement efforts in the acute care and critical access hospital settings. These include acute myocardial infarction (AMI), atrial fibrillation, heart failure, pneumonia, stroke, and surgical infection prevention. As a quality improvement organization, VHQC has the challenge to improve rates on quality indicators in each of these areas.

Geographical analyses are used to identify those providers and provider areas most in need of improvement. Several different types of analyses are used. The examples that follow detail these analyses and provide examples of the output produced.

EXAMPLE ONE

In this first example, the geographical coordinates of Virginia acute care hospitals have been entered into a SAS data set. This data set then can be combined with other data sets containing hospital information as well as rates for selected clinical indicators. This type of graphic is used to identify clusters of hospitals in the state with similar needs. This example map color-codes locations by a rural-urban indicator. Color definitions are as follows:

Black:	Urban
Blue:	Rural
Red:	Rural Critical Access

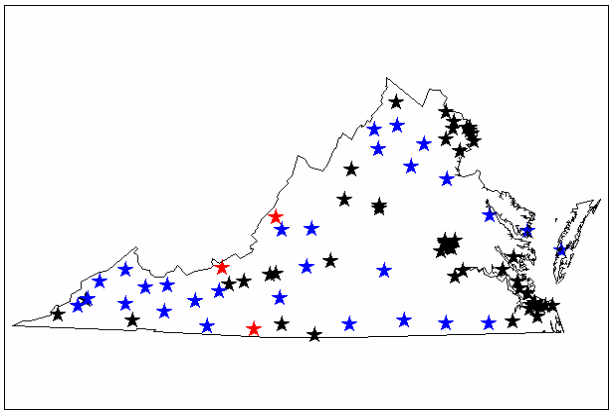
SAS Code:

```
/* Set the graphics environment */
goptions reset=global gunit=pct cback=white
        colors=(black blue green red)
        border htitle=6 htext=3;
pattern value=mempty repeat=49 color=black;

/* Create Annotate data set */
data providers;
   set p.provider(keep=lat long f6);
   lon=-1*long;
   length function style color $8 position $1
         text $20;
   retain function 'label' xsys ysys '2'
         hsys '1' when 'b' state 100;
   state+1; color=f6; size=7; text='V';
         position='5'; style='marker';
   x=lon*arcos(-1)/180; y=lat*arcos(-1)/180;
   output; run;

/*Create Virginia data set */
data virginia;
   set maps.states;
   if state=51; if density <4;
run;

/*Combine state with healthcare providers */
data all; set virginia providers;
run;

/* Project the ALL data set */
proc gproject data=all out=allp project=albers;
   id state; run;

/*Separate projected map & annotate data sets*/
data providersp virginiap;
   set allp;
   if state>100 then output providersp;
   else output virginiap;
run;

/* Show the annotated map */
proc gmap data=virginiap map=virginiap all;
   id state;
   choro state / nolegend midpoints=old
                 annotate=providersp;
run; quit;
```

The ALBERS equal area projection method was used. Due to the algorithms of the projection method, it was necessary to merge the annotate data set with the base map and project the combined data set.

Because SAS requires the use of an ID variable for mapping, the mapping ID variable state was added to the annotate data set. An initial value of 100 was given to not conflict with any value of the state variable in the SAS states data set.

Example One Graphic:



EXAMPLE TWO

In this example, hospital healthcare results are mapped by region. Regional rates are computed on a regular basis and distributed to the providers, along with the overall state rates and their individual hospital rates. This provides the providers with a look at how they performed within their geographical region as well as their overall performance.

Because meetings with providers are held in each region, mapping overall regional results for clinical indicators is also important for identifying strengths and weaknesses within a region and therefore identifying talking points and focus areas specific to a region. Successes in regional rates are also publicized through regional print and broadcast media.

Regions as depicted in the example below are regions as designated by Virginia Health Information (VHI), a non-profit organization providing health information products, for hospital and nursing home health information geographical groupings. Results as mapped by VHQC in this example are for acute care hospitals.

SAS Code:

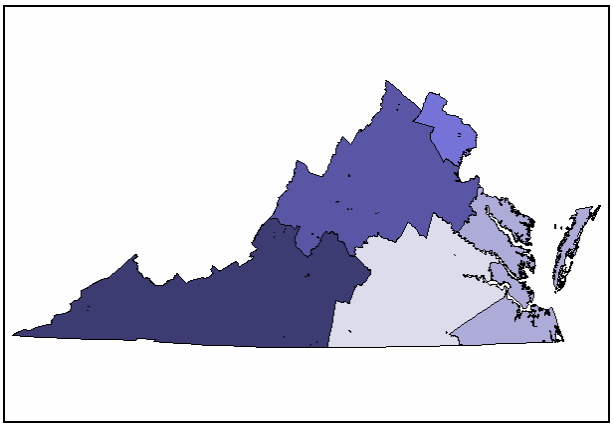
```
/* Set the Graphics Environment */
goptions reset=global gunit=pct colors=(bwh
   vpab vlib lib mob) border;run;

/*Read in regional data set */
data vacntyreg(drop=st_fips);
  length county 5;
  set vacntyregion(drop=county state);
  cnty=left(cnty_fips);
  county=substr(cnty,3,3);
  state=st_fips; run;

/* Create and project Virginia county map */
data vacntymap; set maps.counties;
  if state=51; run;
proc gproject data=vacntymap out=virginia
  project=albers; id county; run;
```

```
/* Merge county and regional information */
data vareg;
  merge virginia vacntyreg;
  by state county;
run;

/* Remove county boundaries */
proc sort data=vareg; by region;run;
proc gremove data=vareg out=varegion;
  by region; id county; run;

/* Produce regional choropleth map */
proc gmap map=varegion data=varegion all;
id region; choro region/nolegend
   coutline=black;
run;
quit;
```

The unprojected county map is used and then projected with the ALBERS projection method. County boundaries are removed with the GREMOVE procedure. Region replaces county as the geographic ID variable. Degrees of a single color are used to display results, with the darker colors identifying the highest rates.

Although this technique is used to identify regional rate differences, the graphic below is an example only and does not reflect any real rate differences.

Example Two Graphic:



EXAMPLE THREE

In this example, each clinical topic is compared regionally by percentage of acute care hospitals showing improvement on a majority of the indicators for that topic. In the example topic, there are 8 indicators for the topic. A hospital would need to show improvement in 5 of the 8 to be tallied as “Improving” and decline in 5 of the 8 to be tallied as “Declining”. Otherwise, they are tallied as “Neither.” Additional regional comparisons are also produced for each individual indicator within a topic.

This type of graphic enhances rate information by providing graphical detail of provider contribution to the regional rates.

Data as depicted in this example are for illustration only and are not intended to depict any actual clinical results.

SAS Code:

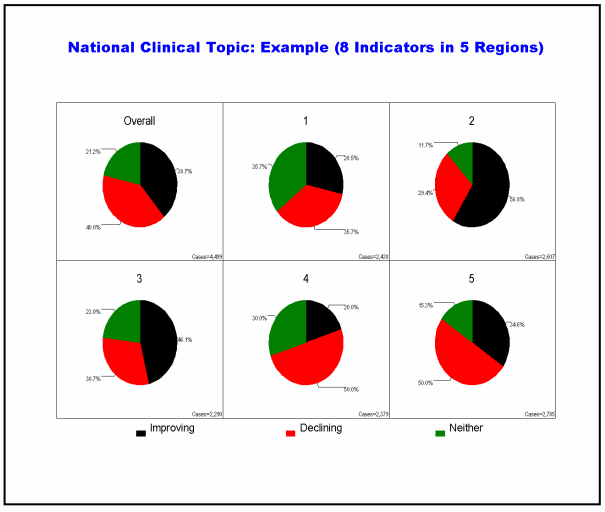
```
/* Set graphics environment and options*/
goptions reset=global gunit=pct device=bmp
   targetdevice=bmp ctext=black graphrc
   interpol=join htext=1.8 vsize=1 in
   nodisplay;  options nobyline;
```

```
/* Set format */
proc format;
   picture pctfmt 0-high='000.0%';

/* Create legend as annotate data set */
data legend;
   length text $ 26 function $ 8 style $ 13
   xsys ysys $ 1; when ="A"; position ='6';
   input function $ x y color $ xsys $ ysys $
      text $ 26.;
   if function='BAR' then do;
      line=3; style='SOLID'; end;
   else do; style='"ARIAL"'; end;
cards;
move 16 65 . 2 1 .
bar 1.8 0.8 black A A .
Move 1 0 1 A A .
CNTL2TXT . . . . . .
Label . . Black . . More Improving
move 46 65 . 2 1 .
bar 1.8 0.8 red A A .
Move 1 0 1 A A .
CNTL2TXT . . . . . .
Label . . Black . . More Declining
move 76 65 . 2 1 .
bar 1.8 0.8 green A A .
Move 1 0 1 A A .
CNTL2TXT . . . . . .
Label . . Black . . Equal
; Run;

/* Set legend graphics space */
proc gslide name="gslide2" annotate=legend;
title1 " ";title2 " ";
note " ";
run;

/* Read in and prepare clinical data set */
data one; set clinical(rename=value=newval);
   length value $26 ;format value $26.;
   retain topnum 0;
   if topic ne lag(topic) then topnum+1;
   format percent 5.1 cnt3sum comma5.;
run;

/* Create macro for pie charts by topic */
%macro pie;
%do i = 1 %to 6;
  data two;
     set one(where=(topnum=&i and category=1));
     call symput("topic",topic);run;
  proc sort data=two; by region cnt3sum; run;
  proc gslide name="gslide1";
     title1 j=c h=3.8 c=blue f="arial black"
     "National Clinical Topic: Example";
     note " ";
  run; quit;
  goptions vsize=10 in  hsize= 12 in htext=6
     ftext="arial narrow"; title1;
  title2 h=10 c=black f="arial narrow"
"#byval(region)";
  footnote2 j=r c=black h=6 f="arial narrow"
     "Cases=#byval(cnt3sum)";
  pattern1 value=SOLID;
  run;
  proc gchart data=two;
     by region cnt3sum;
     Pie value /clockwise
        sumvar=percent noheading  value=arrow
           slice=none midpoints="Improving"
              "Declining" "Neither";
     format percent pctfmt. region $regfmt. ;
  run; quit;
%end; %mend; %pie;
```

```
/* Create template to display pie charts */
goptions gunit=pct;
proc greplay igout=gseg nofs tc=tempcat;
tdef inpadv
1/llx=10 lly=80 ulx=10 uly=90 urx=91 ury=90 lrx=91 lry=80
2/llx=10 lly=10 ulx=10 uly=20 urx=91 ury=20 lrx=91 lry=10
3/llx=10 lly=50 ulx=10 uly=80 urx=37 ury=80 lrx=37 lry=50
   color=black
4/llx=37 lly=50 ulx=37 uly=80 urx=64 ury=80 lrx=64 lry=50
   color=black
5/llx=64 lly=50 ulx=64 uly=80 urx=91 ury=80 lrx=91 lry=50
   color=black
6/llx=10 lly=20 ulx=10 uly=50 urx=37 ury=50 lrx=37 lry=20
   color=black
7/llx=37 lly=20 ulx=37 uly=50 urx=64 ury=50 lrx=64 lry=20
   color=black
8/llx=64 lly=20 ulx=64 uly=50 urx=91 ury=50 lrx=91 lry=20
   color=black;

/* Write results to bitmap file */
goptions display hsize=14 in vsize=12 in;
Template inpadv;
Treplay
   1:gslide1 2:gslide2 3:gchart4  4:gchart
   5:gchart1 6:gchart2 7:gchart3  8:gchart5;
run;
```

The pie charts use the arrow option to print non-overlapping values. The number of cases meeting the criteria for the clinical indicator is included as a footnote in each pie chart.

Example Three Graphic:



EXAMPLE FOUR

Combination plot and vertical bar charts are used to compare changes in regional and individual hospital rates with other factors including number of contacts, number and type of interventions, demographics, and patient mix factors. This information is useful, not only for the hospitals to compare their progress, but also for VHQC to identify and explore any differences. In this example, clinical indicator rates for baseline and four subsequent time periods are compared with the total number of interventions in place in each region during each of the same time periods.

In the example code that follows, Proc GCHART VBAR is used to create the bar chart and annotate data set is created for the plot line. Since both charts can coexist on a single scale axis, it is possible to use GCHART to chart one of the variables. When a

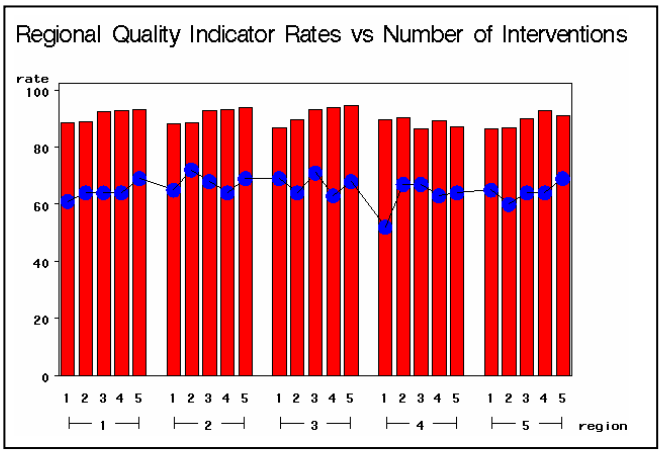
different scale is needed for the second axis or second variable, both the bar chart and the plot are created as annotate data sets. Again, the data depicted in this example is for illustration only and does not depict any actual clinical results.

SAS Code:

```
/* Set Graphics Environment */
goptions reset=global gunit=pct ;run;

/* Read in data set */
Data sample;
  Infile "sample.dat";
  Input region int_time rate interventions;
Run;

/* Create Annotate Data Set */
data anno;
   length function color $8 text $10;
   retain xsys ysys '2' when 'a';
   set sample end=last; by region;
   if _n_ = 1 then function='move';
   else function='draw'; xsys='2';ysys='2';
   x=int_time; y=interventions; group=region;
output;

function='symbol';xsys='2';ysys='2';x=int_time;
   y=interventions; group=region;
   text='dot';size=2;color='blue';output;
run;

/* Create Gchart */
axis1 order =0 to 100 by 20 minor=none;
axis2 label=none;
title "Regional Quality Indicator Rates vs
   Number of Interventions";
proc gchart data=sample annotate=anno;
   vbar int_time/sumvar=rate group=region
   midpoints=1 to 5 by 1 raxis=axis1
maxis=axis2;
run;quit;
```

The Annotate data set used in the output below uses the annotate GROUP variable to match with the GCHART group variable. Because the first default color is red, the annotate symbol is set to Blue to contrast.

Example Four Graphic:



EXAMPLE FIVE

In this example, a report is created which provides information about quality indicator improvement by provider within region. Providers are compared by region with both their regional rate and with the overall state rate. This process was used to identify possible success stories for presentations at regional provider meetings and was followed for each of fifteen inpatient clinical indicators. Hospitals with the most improvement for each clinical indicator with identifiable improvement processes were recruited as speakers for regional meetings across the state.

SAS Code:

```
/* Read in data set */
data rates;
  infile rates;
  input region $ hospital $ rate1 rate2
     rate3 count1 count2 count3;
  pctchange=rate1-rate3;
run;

/* Sort by percent change */
proc sort data=rates; by pctchange;

/* Open ODS output rtf destination and set
   style template  */
ods rtf file="c:\working sas\sesug\report2.rtf"
     style=rtfinp;

/* Eliminate default title */
title;

/* Create the report, grouping by rank within
   region */
proc report data=rates nowd headline pspace=1
      nocenter colwidth=5;
   column region hospital pctchange count1
      rate1 count3 rate3;
   define region /group "Region" order;
   define hospital/ "Hospital";
   define pctchange /noprint order=data;
   define rate1 / mean "Baseline"
      weight=count1;
   define count1/ sum noprint;
   define rate3 / mean "Remeasurement"
      weight=count3;
   define count3/ sum noprint;
   break after region /summarize suppress
      skip style=[font_weight=bold];
   rbreak after/summarize
      style=[font_weight=bold];
run;

/* Close ODS destination */
ods rtf close; run;
```

Because rate1 and rate3 represent different measurement periods with a different sample size, separate weight variables are used for each of these, allowing for weighted means. An ODS template is used that includes font, summary, and title bar definitions. Because the template did not include the bolding of summary lines, this is included in the code.

Example Five Output:

Region	Hospital	Baseline	Remeasurement
1	B	91.9	93.3
	D	94.3	95.4
	A	91.3	91.6
	C	93.3	93.5
	E	88.1	84.5
Mean		***92.31***	***92.98***

Region	Hospital	Baseline	Remeasurement
2	I	88.2	90.2
	F	91.1	92.3
	H	86.1	87.2
	G	100	100
Mean		***89.81***	***91.02***
3	N	88.3	90.8
	J	90	92.3
	L	95.4	97.4
	M	93.2	94
	K	84.2	80
Mean		***91.48***	***92.89***
Overall Mean		***91.25***	***92.37***

OUTPATIENT

Clinical areas have also been identified for by the Centers for Medicare & Medicaid Services for quality improvement efforts in the outpatient setting. These include mammography, flu and pneumoccocal immunizations, dilated eye exams for diabetic patients, lipid profiles for diabetic patients, and HbA1c testing for diabetic patients. The VHQC has the challenge to improve rates on these indicators for the Medicare population in the state of Virginia.

Geographical analyses are used to target high Medicare beneficiary population areas with low indicator rates to maximize the effects of limited resources.

EXAMPLE ONE

In this example, maps are used to target counties with low indicator rates and a high Medicare beneficiary population. Counties with high beneficiary populations and low indicator rates are targeted for social marketing campaigns. Providers who practice in those counties are also contacted as potential collaborators to receive educational materials in quality improvement and standards of practice for the indicator in question.

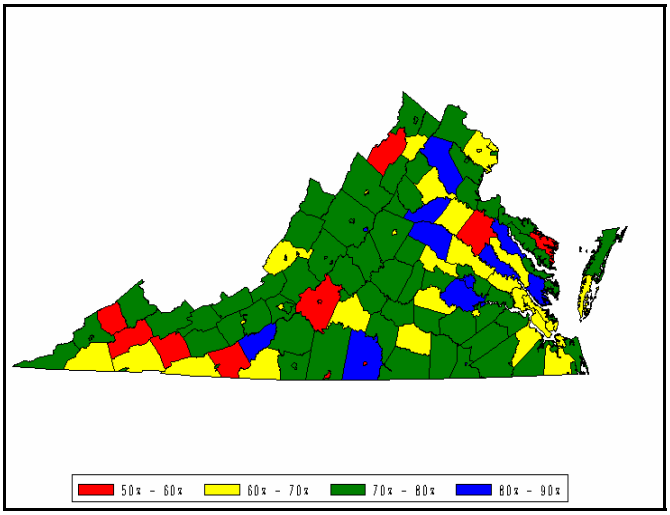
This example shows a map of biennial lipid profiles for counties in the state of Virginia. It uses the GMAP procedure and the GNOMON projection method.

Data used in this example are for illustration purposes only, they do not represent actual county clinical results.

SAS Code:

```
/* Set graphics environment and formats */
goptions rotate=landscape targetdevice=dj1600c ;
run;
proc format;
  value rfmt      50 -< 60 = "50% - 60%"
                  60 -< 70 = "60% - 70%"
                  70 -< 80 = "70% - 80%"
                  80 -< 90 = "80% - 90%";
run;
pattern1 color=red value=solid;
pattern2 color=yellow value=solid;
pattern3 color=green value=solid;
pattern4 color=blue   value=solid;
pattern5 color=cyan value=solid;
pattern6 color=magenta value=solid;

/* Read in county rates */
data County;
   merge interim2.T1county(in=insamp)
         where=(label="totlipd"))
         sdata.ssafipco(in=inc);
   by ssacnty;
   if insamp and inc;
   county=input(fipcode,3.);
run;

/* Create Virginia County Dataset */
data vacnty;
 set maps.counties;
 if state=51;
run;
/* Project Virginia County Map */
proc gproject data=vacnty out=p_vacnty
    project=gnomon;
    id county;
run;

/* Create a Map of Virginia Rates */
proc gmap map=p_vacnty data=allsamps;
   format rate rfmt.;
   id county;
   choro rate/ coutline=black legend=legend
   discrete;
   label rate="Lipid Rate";
   legend value=(j=center) frame label=none;
run;
```

Example One Graphic:



EXAMPLE TWO

In the second example, boxplots are utilized to identify counties or providers whose quality indicator rates fall above or below the population of their peers over time. Counties or providers which fall below their peers can be targeted for additional interventions. Counties or providers which fall above their peers can be a focus for research into best practices, which can then be communicated to the Virginia health provider community.

Boxplots were created using Proc GPLOT, with Interpol=Boxt10 and extra plots to include data points and statewide rates,

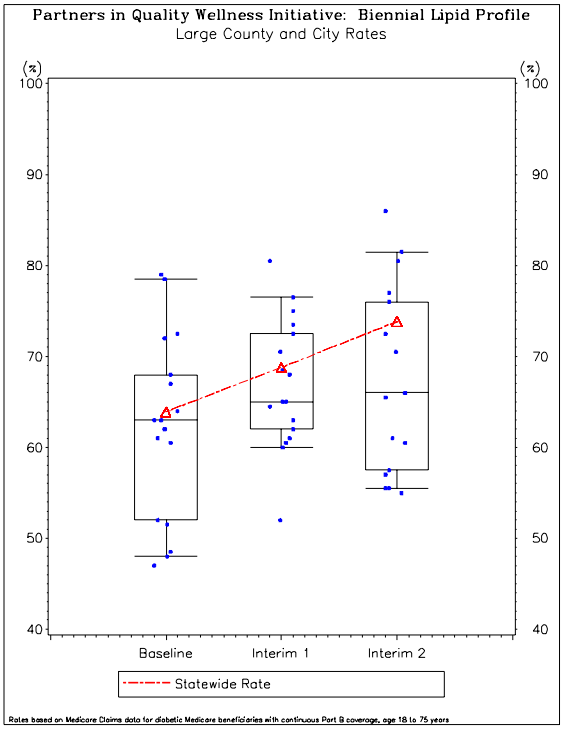
provider specific rates, or a high benchmark rate such as best practice rate.

Data used in this example are for illustration purposes only, they do not represent actual county or statewide clinical results.

SAS Code:

```
/* Read in dataset with County Rates    */
/* Keep counties with 800 benes or more */
data county_all2;
 set county_all;
 by county;
 if cntyname not in (' ','UNKNOWN');
 if count1 ge 800;
 rand=ranuni(234);
 if rand gt .85 then time2=time + .10;
  else if rand gt .75 then time2=time + .07;
   else if rand gt .60 then time2=time + .04;
    else if rand gt .50 then time2=time +.01;
     else if rand gt .40 then time2=time-.01;
      else if rand gt .30 then time2=time-.04;
       else if rand gt .15 then time2=time-.07;
        else time2= time - .10;
    label perc2="Statewide Rate"
          perc1="  ";
run;

/* Macro which creates graph */
%macro
graph(dat,var,min,titl,titl2,bdate,rdate1,
           rdate2,subperc,subset,agegrp);
data _null_;
  set &DAT;
  where label="&var";
run;
proc format;
  value tft 0=" "
            1="Baseline"
            2="Interim 1"
             3="Interim 2"
            4=" ";
run;

/* Set GPLOT options */
goptions reset=global border htitle=1.5
htext=1.3
ftitle=complex ftext=simplex device=sasemf;

/* Set up symbol for Boxplot */
 symbol1 interpol=boxt10
         mode=exclude
         value=none
         co=black
         cv=black
         height=.6
         bwidth=15
         width=5;
/* Set up Symbol for Data Points */
 symbol2 ci=blue
          value=dot
           h=.5;
/* Set up symbol for Statewide Rate */
  symbol3 ci=red
          i=join
          value=triangle
          h=1.7
          line=8
          width=2;
/* Define X-axis */
   axis1 split='\' label=none
       order=(0 to 4 by 1)
       length=100
       ;
/* Define Y-axis */
  axis2 label=(h=1.5 '(%)')

   order=(&MIN to 100 by 10);

/* Define Legend */
  legend1 position=(middle bottom outside)
          frame
          across=2
          label=none
          shape=line(10)
          value=(h=1.3);
 run;

/* Define Output File */
ods rtf file=
"c:\boxplots\LargeCountyPlotFor%CMPRES(&var).rtf";

/* Plot Graph */
proc gplot data=&DAT;
where label="&var";
plot rate*time rate*time2/ overlay haxis=axis1
   vaxis=axis2;
plot2 perc2*time/overlay haxis=axis1
vaxis=axis2 legend=legend1;
   format time tft.;
run;
ods rtf close;
%mend graph;

/* Call MACRO */
%graph(county_all3,totlipd,40,Biennial Lipid
 Profile,Large County and City Rates,11/01/97-
 10/31/99,11/01/98 - 10/31/00,05/01/99 -
 04/30/01,,diabetic,18 to 75 years)
quit;
```

Example 2 Graphic:



Boxplots show the range, median, quartiles, and change of the data over time. Additionally, elements such as mean, statewide rate, and upper limits are added to the boxplots when applicable.

CONCLUSION

This paper demonstrates several methods for using geographical data displays to target areas for health care quality improvement. Geographical targeting provides for more focused social marketing campaigns, allowing for the most efficient use of quality improvement dollars. The built-in flexibility of SAS, ODS, and SAS/GRAPH, as well as the capabilities of SAS Annotate, facilitate the development of data analysis processes that can be used to demonstrate geographical differences and to provide a foundation for further analyses to identify additional relationships and critical factors to enhance the quality improvement process.

REFERENCES

http://www.vhqc.org.

SAS Institute Inc. 1999. *SAS/GRAPH Software: Reference, Version 8, Volumes 1 and 2.*

SAS Institute Inc. 1990. *SAS/GRAPH Software: Reference, Version 6, First Edition, Volumes 1 and 2.*

Virginia Health Information. *Hospitals and Nursing Homes. 1997 Virginia Industry Report*.

ACKNOWLEDGEMENTS

Data sets used for the inpatient examples were generated used SAS/STAT Proc SURVEYSAMPLE.

Virginia hospital locations were geocoded using Tele Atlas at www.geocode.com.

The analyses upon which this publication is based were performed under contract no. 500-99-VA-03, titled “Utilization and Quality Control Peer Review Organization for the Commonwealth of Virginia,” sponsored by the Centers for Medicare & Medicaid Services (CMS), Department of Health and Human Services (DHHS). The content of this publication does not necessarily reflect the view or policies of the DHHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The Virginia Health Quality Center (VHQC) assumes full responsibility for the accuracy and completeness of the ideas presented. This publication is a direct result of the Health Care Quality Improvement Program initiated by CMS, which has encouraged identification of quality improvement projects derived from analysis of patterns of care and therefore required no special funding on the part of VHQC. Ideas and contributions to the VHQC concerning experience in engaging with the issues presented are welcomed.

CONTACT INFORMATION

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

Barbara B. Okerson, Ph.D.
Biostatistician
Virginia Health Quality Center
4510 Cox Road, Suite 400
Glen Allen, VA 23060
Phone: 804.289.5320
e-mail: vapro.bokerson@sdps.org

Charlotte F. Carroll, M.S.
Biostatistician
Virginia Health Quality Center
4510 Cox Road, Suite 400
Glen Allen, VA 23060
Phone: 804.289.5320
e-mail: vapro.ccarroll@sdps.org

VIRGINIA HEALTH QUALITY CENTER