Using SAS for Error Grid Analysis (EGA) of Glycated Hemoglobin A1c
Jesse A. Canchola, Tethys Bioscience, Emeryville, California USA
Christopher M. Canchola, University of California, Santa Cruz, California USA

ABSTRACT
Error Grid Analysis (EGA) uses graphical error grids for performance evaluations of medical devices* or for comparing candidate and comparative measurement methodologies**. Though not as widely used as we would like, EGA is an instructive and visually appealing performance evaluation methodology that shows the measurement of a new or improved test on the y-axis with a paired measurement of the reference test on the x-axis overlaid on a colorful grid of clinical regions (as established by experienced clinicians by consensus). Although error grids for glucose (mg/dL and mmol/l) and hemoglobin (g/dL) have been proposed, there is no known EGA for glycated hemoglobin A1c (HbA1c) (%) using The SAS® System v9.3, including SAS/GRAPH® and the SAS/GRAPH® Annotate facility.

* For example, a new device as compared to a predicate or currently accepted or gold-standard device.
** For example, a new blood collection method as compared to a current standard.

INTRODUCTION
EGA was developed by Clarke and colleagues (1987; Clarke, et al.) originally for clinical performance evaluation of blood glucose values as obtained from a blood glucose meter and compared to a reference value as the gold standard – as opposed to the traditional statistical method comparison approaches such as Deming regression, Passing-Bablok regression or Ordinary Least Squares regression (OLS). Subsequently, the phrase, “Clarke Error Grid” was coined to describe this type of analysis for glucose and, eventually, non-glucose analyte comparisons of this kind were simply called, Error Grid Analysis or EGA.

Typically, EGA consists of a pre-defined grid of several regions that indicate whether the measured test vs. reference data paired observations are within clinically acceptable boundaries. As an example, Figure 1 shows a typical five-region Clarke Error Grid of glucose measurements (mg/dL) for diagnosis and treatment of hypo- and hyperglycemia from a blood sample drawn by fingerstick versus the reference sample from a venous draw. The five regions are coded such that:

- Region A (light green) should contain values that are within 20% of the reference value and are considered, "clinically accurate";
- Region B (yellow) will contain values that are greater than 20% of the reference value but would "not lead to inappropriate treatment";
- Region C (orange) contains values that would "lead to unnecessary treatment";
- Region D (pink) contains values that would indicate a "potentially dangerous failure to detect hypo- or hyper-glycemia";
- Region E (red) contains values that would confuse treatment of hypoglycemia for hyperglycemia and vice-versa.

To our knowledge, at this time, there is no known EGA for glycated hemoglobin A1c in percent units. Therefore, a proposed EGA solution for glycated hemoglobin A1c (in %) is presented.
DEVELOPMENT

The motivation for developing an HbA1c Error Grid comes from using HbA1c as a method for monitoring the degree of glucose metabolism or regulation. Table 1 shows the American Diabetes Association’s (ADA) classification of glucose regulation by HbA1c for patients with "normal glucose regulation," "prediabetes," and "diabetes".

<table>
<thead>
<tr>
<th>HbA1c (%)</th>
<th>Normal Glucose Regulation</th>
<th>Prediabetes</th>
<th>Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 5.7 %</td>
<td></td>
<td>5.7 % to less than 6.5%</td>
<td>Greater than or equal to 6.5 %</td>
</tr>
<tr>
<td>Greater than or equal to 6.5 %</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Classification of glucose regulation using HbA1c (ADA, 2013)

From this classification along with analogues to the Clarke Error Grid methodology (Morey et al., 2011), it is straightforward to develop a simple grid with three categories:

- Region A (light green) will contain values that are within 20% of the reference value and are considered, "clinically accurate";
- Region B (yellow) will contain values that are greater than 20% of the reference value but would "not lead to inappropriate treatment";
- Region C (red) contains values that would confuse treatment decision for diabetes.

The SAS® System v9.3, including SAS/GRAPH® and the SAS/GRAPH® Annotate facility are used to create the three regions. A SAS/Graph Annotate dataset is created that defines the points (for connecting lines), labels, and colors for each region.

For example, the first region, Region A, is created or "drawn" with the following SAS/Graph Annotate dataset code:

```sas
data anno_colored_areas ;
length function color style $8 ;
xsys='2' ; ysys='2' ; hsys='3' ; when='b' ; color='black' ; size=.1 ;
* green area ;
color="cx62ff00" ; style='solid' ; function='poly' ; x=4 ; y=4 ; output ;
function='polycont' ;
x=5.7 ; y=4.0 ; output ; x=5.7 ; y=5.4 ; output ;
x=7.2 ; y=6.5 ; output ; x=12.0 ; y=6.5 ; output ;
x=12.0 ; y=12.0 ; output ; x=6.5 ; y=12.0 ; output ;
x=6.5 ; y=7.2 ; output ; x=5.4 ; y=5.7 ; output ;
x=4.0 ; y=5.7 ; output ; x=4.0 ; y=4.0 ; output ;
```

and so on for the remaining regions (see Appendix A). Two important functions in SAS/Graph Annotate facility are the POLY and the POLYCONT. The POLY function (shown above as, function='poly'), specifies the beginning of a polygon. The POLYCONT function (shown as function='polycont' in the code above) specifies successive points (shown in the code above as x=## ; y=##) for connecting a continuous polygon begun at the point identified by the POLY function (in this case, x=4 ; y=4).

After the SAS/Graph Annotate dataset is created, the necessary PROC GPLOT options are:

```sas
proc gplot data=A1C_data anno=anno_all ;
plot FS_hbA1c*Ven_hbA1c=1 FS_hbA1c*Ven_hbA1c=2 ; run ;
```

where "anno_all" in the code "anno=anno_all", above, is the combined set of defined SAS/Graph Annotate datasets (see Appendix A).

Figure 2 shows the results of applying the complete SAS code (Appendix A), including PROC GPLOT, for the three regions. The sample observations shown are the Venous vs. Fingerstick data pairs (shown as blue dots) with an OLS regression line in solid blue following the 1-to-1 dotted unity line.
The reader is encouraged to explore extensive coverage of the SAS/Graph Annotate facility as given elsewhere (for example: Carpenter, 2006; Mink and Pasta, 2006 and others). In addition, there are different methods for constructing error grids. The document, "EP27-A, How to Construct and Interpret an Error Grid for Quantitative Diagnostic Assays; Approved Guideline" from the Clinical and Laboratory Standards Institute (CLSI; 2012) gives a brief history of EGA and proposes methods, based on clinical consensus, for constructing error grids that can be applied to different areas of study.

**CONCLUSION**

The authors endeavored to show a unique EGA method for hemoglobin A1c as a percent. The error grid regions were determined using ADA guidelines and created using The SAS System v9.3, including the SAS/GRAPH® and the SAS/GRAPH® Annotate facilities, with sample data overlaid as an example.

**REFERENCES**


CONTACT INFORMATION

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

Jesse A. Canchola
667 Alquire Parkway
Hayward, CA 94544
E-mail: JCanch@Gmail.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.

Disclaimer: The views and opinions expressed in this article are solely those of the authors and do not necessarily reflect the official policy or position of Tethys Bioscience, Inc.

APPENDIX A. SAS CODE FOR HBA1C ERROR GRID

* "Using SAS for Error Grid Analysis (EGA) of Glycated Hemoglobin A1c";
* WUSS 2014 - Planet Hollywood, Las Vegas, NV;
* November 13 to 15, 2013;
%let name=hemoglobin_error_grid;
filename odsout '.';
%let dataclr=cx4da5f7;
%let grayclr=gray33;

* mock A1c data;
data A1C_data;
label FS_hbA1c='Fingerstick HbA1c (%)';
label Ven_hbA1c='Venous HbA1c (%)';
input Ven_hbA1c1 FS_hbA1c1;
FS_hbA1c = FS_hbA1c1;
Ven_hbA1c = Ven_hbA1c1;
datalines;
< DATA EDITED OUT FOR SPACE PURPOSES: CONTACT AUTHOR >;
run;

* set up the html hover-text for the red plot marker;
data A1C_data; set A1C_data;
length hover_text $100;
hover_text='title='||quote('FS_hbA1c (%) = '||trim(left(FS_hbA1c))||'0d'||'Ven_hbA1c (%) = '||trim(left(Ven_hbA1c)))||'
href="hemoglobin_error_grid_info.htm"';
run;

* set up each colored grid section;
data anno_colored_areas;
length function color style $8;
xsys='2'; ysys='2';
hsys='3'; when='b';
color='black'; size=.1;

* green area;
color='cx62ff00'; style='solid';
function='poly';
x=4; y=4; output;
function='polycont';
x=5.7 ; y=4.0 ; output;
x=5.7 ; y=5.4 ; output;
x=7.2 ; y=6.5 ; output;
x=12.0 ; y=6.5 ; output;
x=12.0 ; y=12.0 ; output;
x=6.5 ; y=12.0 ; output;
x=6.5 ; y=7.2 ; output;
x=5.4 ; y=5.7 ; output;
x=4.0 ; y=5.7 ; output;
x=4.0 ; y=4.0 ; output;

* bottom yellow area;
color="yellow"; style='solid';
function='poly';
x=5.7 ; y=4.0 ; output;
function='polycont';
x=5.7 ; y=5.4 ; output;
x=7.2 ; y=6.5 ; output;
x=12.0 ; y=6.5 ; output;
x=12.0 ; y=5.7 ; output;
x=6.5 ; y=5.7 ; output;
x=6.5 ; y=4.0 ; output;
x=5.7 ; y=4.0 ; output;

* top yellow area;
color="yellow"; style='solid';
function='poly';
x=4.0 ; y=5.7 ; output;
function='polycont';
x=4.0 ; y=5.7 ; output;
x=6.5 ; y=7.2 ; output;
x=6.5 ; y=12.0 ; output;
x=5.7 ; y=12.0 ; output;
x=5.7 ; y=6.5 ; output;
x=4.0 ; y=6.5 ; output;
x=4.0 ; y=5.7 ; output;

* top red area;
color="cex5252"; style='solid';
function='poly';
x=4.0 ; y=6.5 ; output;
function='polycont';
x=4.0 ; y=12.0 ; output;
x=5.7 ; y=12.0 ; output;
x=5.7 ; y=6.5 ; output;
x=4.0 ; y=6.5 ; output;

* bottom red area;
color="cex5252"; style='solid';
function='poly';
x=6.5 ; y=4.0 ; output;
function='polycont';
x=6.5 ; y=5.7 ; output;
x=12.0 ; y=5.7 ; output;
x=12.0 ; y=4.0 ; output;
x=6.5 ; y=4.0 ; output;
run;

data anno_outline; set anno_colored_areas;
style='empty'; color="&grayclr";
run;

data anno_diagonal;
length function color $8;
xsys='2'; ysys='2';
hsys='3'; when='b';
color="black"; size=.1;
function='move'; x=4; y=4; output;
function='draw'; x=12.0; y=12.0; line=33; output;
run;

data anno_letters;
length function color $8;
xsys='2'; ysys='2';
hsys='3'; when='b';
color="&grayclr"; size=.;
function='label'; position='5';
text='A'; x=10; y=9; output;
text='B'; x=6.1; y=10; output;
text='C'; x=5; y=10; output;
text='A'; x=9; y=10; output;
text='B'; x=10; y=6.1; output;
text='C'; x=10; y=5; output;
run;

data anno_all; set anno_colored_areas anno_outline anno_diagonal
anno_letters;
run;

gooptions device=png ypixels=625 xpixels=600;
gooptions noborder;
opts listing close;
opts html path=odsout body="&name..htm"
  (title="Error Grid for PreDx Venous vs. Fingerstick HbA1c (%)")
  style=sasweb;
gooptions gunit=pct htitle=3.5 ftitle="albany amt/bold" htext=2.5
  ftext="albany amt/bold"
axis1 order=(4 to 12 by 2) minor=(number=1) offset=(0,0) length=5.0in
  label=(angle=90);
axis2 order=(4 to 12 by 2) minor=(number=1) offset=(0,0) length=5.0in
  label=(angle=0);
symbol1 value=dot height=2.5 interpol=rl width=2 color=&dataclr
  ci=&dataclr;

* draw a smooth circle around the blue dot *;
symbol2 value=circle height=2.5 interpol=none color=&grayclr;
title1 "";
proc gplot data=A1C_data anno=anno_all;
plot FS_hbA1c*Ven_hbA1c=1 FS_hbA1c*Ven_hbA1c=2 / overlay
  vaxis=axis1 haxis=axis2 noframe
html=hover_text
des='' name="&name"
footnote1 "Figure 2. Error Grid for Venous vs. Fingerstick HbA1c (%)";
run ;
quit ;
ods html close ;
ods listing ;

□