Leveraging RAREEVENTS Procedure Options to Monitor and Evaluate Infrequent Events in Healthcare

Xiaohong (Lynn) Liu, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL
Roderick Jones, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

ABSTRACT

In healthcare, the purpose of statistical process control (SPC) is often to quantify improvements and identify unintended consequences resulting from an intentional change in an environment, policy, treatment protocol, or decision-support tool. Appropriate for measurement of infrequent or low-probability events, the RAREEVENTS procedure is valuable for healthcare quality monitoring and evaluation. The RAREEVENTS procedure was enhanced in the SAS/QC® version 15.1 release and now allows users to apply tests to detect special cause variation. We provide an example of leveraging new PROC RAREEVENTS options, including READPHASE=, PHASELEGEND, LIMITPHASES, TEST1, TEST2, TEST3, TEST4 and TESTACROSS.

INTRODUCTION

The Institute for Healthcare Improvement works with health systems, countries, and other organizations on improving safety, quality, and value in health care (Institute for Healthcare Improvement 2019). One aspect of its endorsement of the science of improvement is the use of statistical process control (SPC) to quantify improvements and identify unintended consequences resulting from an intentional change in an environment, policy, treatment protocol, or decision-support tool. The SHEWHART procedure enables SAS users to generate many SPC charts, but it requires the aggregation and summarization of data by temporal units. To monitor and evaluate an adverse, unexpected or rare event in the healthcare context (e.g., infections associated with device implantation, organ failure following transplant surgery) acquisition of sufficient data is too delayed for SHEWHART procedure charts, and instead a g (Benneyan 2001) or t chart (Santiago and Smith 2013) is preferable.

The RAREEVENTS procedure creates a rare events chart – a plot of the amount of time or number of opportunities between events; an increase in the time or number of opportunities suggests an improvement (Ransdell 2016). SAS/QC® version 15.1 delivers enhanced statistical and visualization features for the RAREEVENTS procedure through the addition of special cause tests, and the READPHASE=, PHASELEGEND, LIMITPHASES, and TESTACROSS options.

This paper describes the foundational syntax for creating a rare events chart of adverse event occurrences in a healthcare context using the RAREEVENTS procedure's CHART statement, as well as the prerequisite data preparation. We share our experiences with defining baseline periods, and we highlight the problems and solutions involved in freezing and extending rare events chart central lines across phases, conducting special cause tests, and identifying and displaying shifts using several PROC RAREEVENTS options. With this orientation, healthcare statisticians have a rare events-specific framework for applying this measurement tool to quality improvement, patient safety and cost-effectiveness initiatives.

GETTING STARTED WITH RARE EVENTS CHARTS

The analysis of rare events makes use of a data set of patients or visits at risk of developing the complication of interest. Our simulation data set, ELIGIBLE, contains a unique identifier for the visit, the date, and a Yes/No value for the variable Event, indicating whether the adverse event occurred in association with the visit (Output 1).

```
                 Visit
Obs      Date     ID     Event

  1    07/06/13   0AEC    Yes
  2    07/10/13   7NQQ    No
  3    07/19/13   7MJQ    No
  4    07/21/13   2HAA    No
  5    07/25/13   0HAH    No
```

Output 1. Excerpt of the Detail-level ELIGIBLE Data Set from a PROC PRINT Statement

PREPARING DATA

Our PROC RAREEVENTS example evaluates the number of opportunities between events. The steps below are required to count the visits between the events from the detail-level data set ELIGIBLE and prepare the date for processing. We ensure the visits are sorted in chronological order, use the automatic variable _N_ to assign the visits a consecutive visit number, and use the LAG function to define the number of visits between events (Output 2).

```sas
%LET Dataset=ELIGIBLE;
%LET StartEvaluation=01Jul13;
%LET EndEvaluation=31May19;

PROC SORT DATA=&Dataset;
     by Date;
run;

DATA &Dataset (WHERE=(Date GE "&StartEvaluation"D AND Date LE
"&EndEvaluation"D));
     SET &Dataset;
     VisitNumber=_N_;
RUN;

DATA EVENTS;
     SET &Dataset (WHERE=(event=" Yes"));
     VisitsBetween=VisitNumber-LAG(VisitNumber);
     LABEL VisitsBetween ="Eligible visits between events";
     KEEP VisitID Date VisitNumber VisitsBetween;
RUN;
```

```
                 Visit    Visit    Visits
Obs      Date     ID     Number   Between

  1    07/06/13   0AEC      1        .
  2    08/13/13   7JJJ      8        7
  3    08/24/13   5QRR     12        4
  4    09/26/13   8JJZ     20        8
  5    10/09/13   0IFI     25        5
```

Output 2. Excerpt of the EVENTS Data Set from a PROC PRINT Statement.

PRODUCING THE RARE EVENTS CHART

The core syntax used to generate the rare events chart is below. The DIST option specifies the probability distribution used (Kaminsky et al. 1992, Benneyan 2001b).

```sas
PROC RAREEVENTS DATA=EVENTS;
     CHART VisitsBetween*Date/
```

```
                DIST=GEOMETRIC;
RUN;
```

To produce the chart, Output Delivery System (ODS) Graphics must be used, so the code above should be enveloped as shown below. (For the remainder of the paper, this ODS GRAPHICS code is implied and will not be repeated.)

```
%LET PATH=C:\Figures;
GOPTIONS RESET=goptions;
ODS GRAPHICS ON/
      IMAGEFMT=PNG  RESET IMAGENAME="Figure 1";
ODS LISTING GPATH="&path." IMAGE_DPI=250;

PROC RAREEVENTS DATA=EVENTS;
      CHART VisitsBetween*Date/
            DIST=GEOMETRIC;
RUN;

ODS GRAPHICS OFF;
```

In the resulting chart, the number of visits (or contacts at risk for an adverse event) between the events are plotted against the visit dates associated with event occurrences (Figure 1). The probability of two events occurring in association with consecutive visits and the probability of 124 visits occurring between events is shown below the chart, along with the median, upper and lower probability limits on the right-side Y-axis.
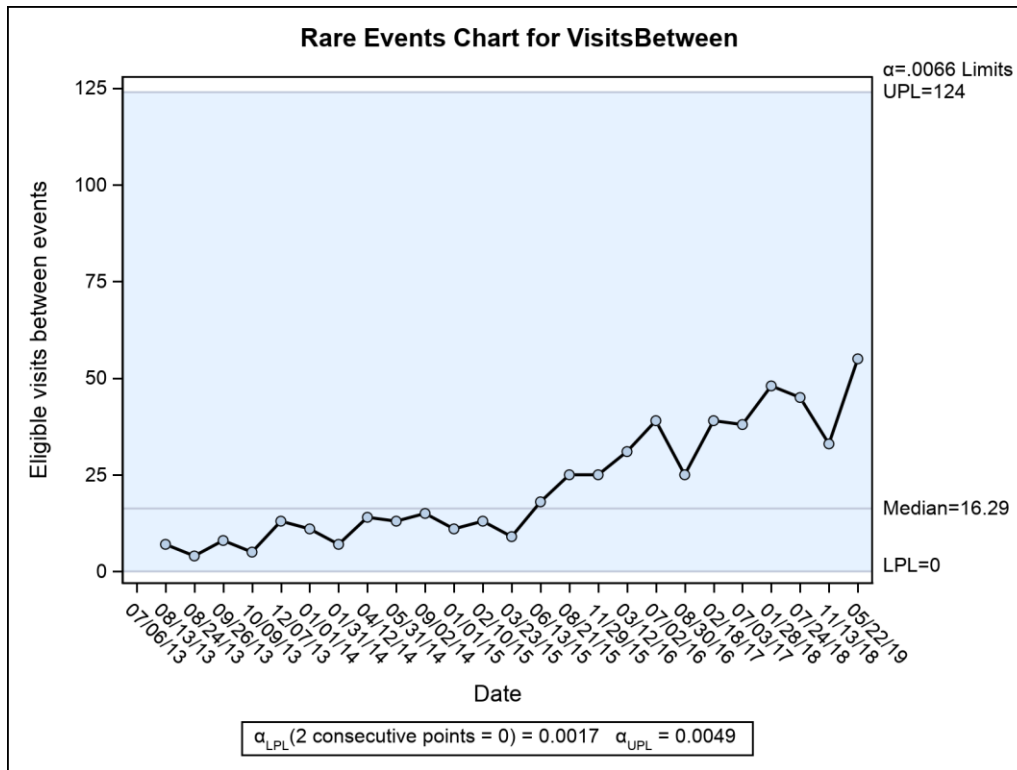

Figure 1. Rare Events Chart.

The upward trend appearing to begin with the June 13, 2015 event suggests that improvement is occurring. RAREEVENTS procedure options must be leveraged to generate statistics to support or refute this impression.

DEFINING A BASELINE

For an approach to detecting improvement through a rare events chart, we consulted *The Health Care Data Guide: Learning from Data for Improvement* (Provost and Murray 2011). The section in Chapter 4 named "Establishing and Revising Limits for Shewhart Charts" highlights several nuances in analyzing the impact of an intervention. These include defining a baseline period and detecting improvements after freezing and extending an initial mean and control limits. The same concepts can apply to rare events charts. A section in chapter 7 titled "Shewhart Charts for Rare Event" provides a formula to calculate the theoretical median for a *g* chart and make this the center line above or below which eight consecutive points can represent a shift.

Provost and Murray propose establishing a baseline center line and control limits based upon at least 20 data points during a time period in which the process of interest was stable. While for other SPC chart types we can count 20 temporal units back from the date of an intervention, for a rare events chart it is likely necessary to apply this guideline in a dynamic way. For instance, instead of counting temporal units, a baseline can be generated by using the 20 events that occurred leading up to the date of the intervention. If events are so rare that 20 pre-intervention points are not available, a smaller number might be used if it represents a time period of process stability.

The data set BASELINE was created to contain only the first 21 events from the EVENTS data set, all of which occurred prior to the end of the baseline period that concluded March 30, 2017. To generate a centerline and limits, we add the OUTLIMITS option to the RAREEVENTS procedure. The NOCHART option suppresses graphic output.

```
PROC RAREEVENTS DATA=BASELINE;
     CHART VisitsBetween*Date/
          DIST=GEOMETRIC
          OUTLIMITS=LIMITS_BASELINE
          NOCHART;
RUN;
```

ADDING PHASES TO RARE EVENTS CHARTS

After distinguishing events associated with the baseline period through the created variable _PHASE_ in the EVENTS data set, two options are added to the RAREEVENTS procedure: READPHASES=ALL specifies the phases to be recognized and PHASELEGEND displays phase labels in a legend across the top of the chart. The OUTLIMITS option creates a data set called LIMITS containing the statistics for each phase.

```
PROC SQL;
     CREATE TABLE EVENTS AS SELECT EVENTS.*,
          CASE
               WHEN VisitID in (SELECT VisitID from BASELINE) THEN
'Baseline'
               ELSE 'Post-baseline'
          END
     AS _PHASE_
          FROM EVENTS;
QUIT;
PROC RAREEVENTS DATA=EVENTS;
     CHART VisitsBetween*Date/
          DIST=GEOMETRIC
          OUTLIMITS=LIMITS
          READPHASES=ALL
          PHASELEGEND;
RUN;
```
In the resulting chart, the two phases have different center lines and probability limits (Figure 2).
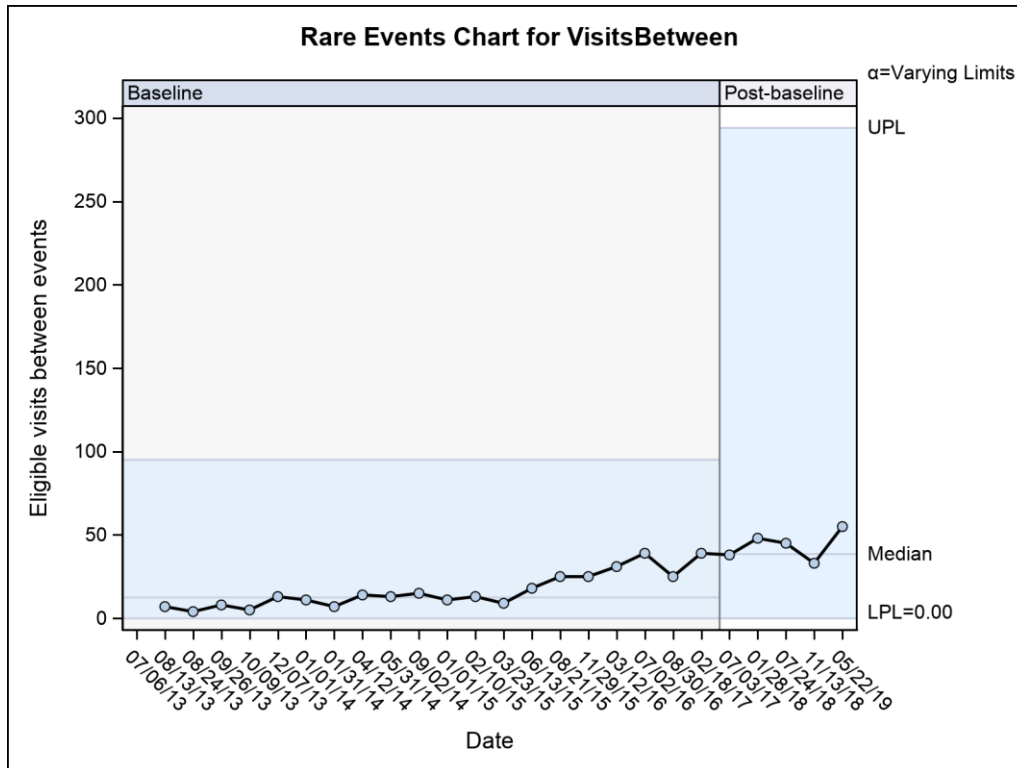
Figure 2. Rare Events Chart with Phases Defined.

EXTENDING BASELINE AND TESTING FOR IMPROVEMENT

With the phase-specific results, we now have a baseline period with which improvements or unintended consequences of the intervention can be detected and statistically quantified.

To "freeze" and extend the baseline for use in testing against the post-baseline, we modify the limits of the chart from the two phases by replacing the post-baseline center line and limits with those of the baseline in a new data set, LIMITS_EXTENDED.

```
PROC SQL noprint;
      SELECT _median_ INTO: centerline
            FROM LIMITS WHERE _phase_="Baseline";

            SELECT _LPL_ INTO: lpl
                  FROM LIMITS WHERE _phase_="Baseline";

                  SELECT _UPL_ INTO: upl
                        FROM LIMITS WHERE _phase_="Baseline";
QUIT;
DATA LIMITS_EXTENDED;
      SET LIMITS;

      IF _phase_="Post-baseline" THEN
            DO;
                  _median_=&centerline;
                  _LPL_=&lpl;
                  _UPL_=&upl;
            END;
RUN;
```

5

Using the statistics from this new dataset and the option LIMITPHASES, we produce a chart with an extended baseline's center line and limits. Because the distribution is established by the LIMITS= data set, the DIST= option is no longer used.

```
PROC RAREEVENTS DATA=EVENTS LIMITS=LIMITS_EXTENDED;
      CHART VisitsBetween*Date/
            READPHASES=ALL
            PHASELEGEND
            LIMITPHASES=ALL;
RUN;
```

Figure 3 provides further visual evidence that improvement occurred, with the number of visits between events consistently above the center line starting with the June 13, 2015 event.
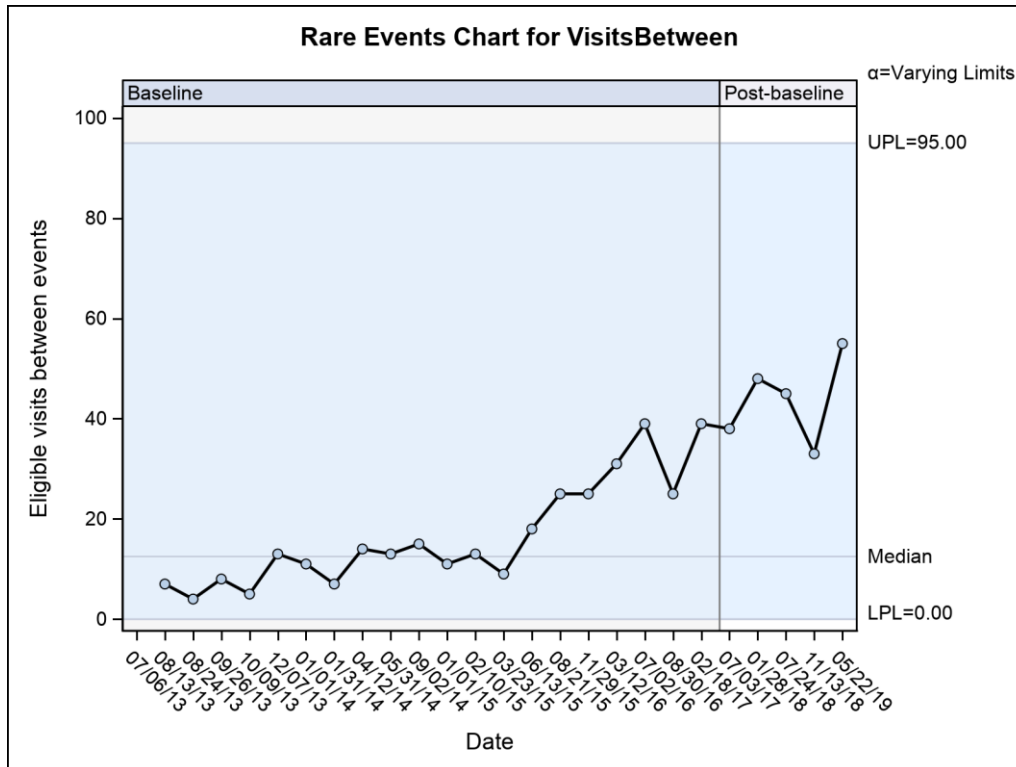
Figure 3. Rare Events Chart with Phases Defined and Baseline Extended.

Tests can now be applied to provide further evidence of an improvement.

TESTING FOR SPECIAL CAUSE VARIATION

PROC RAREEVENTS offers four tests for special cause variation. TEST1 detects one point outside of the probability limits; TEST2 detects a sequence of n or fewer points within which k points are on one side of the center line; TEST3 detects a sequence of consecutive points (default=6) that are steadily increasing or steadily decreasing; TEST4 detects a sequence of consecutive points (default=14) that alternate up and down.

TEST2 aids quality improvement assessments by detecting a process shift in the desired or undesired direction. The number of consecutive *k* and *n* points is adjustable, and an 8-point run is commonly used in healthcare. To apply these four tests to the simulation, it is necessary to add options to the CHART statement: TEST1 TEST2 (K=8 N=8) TEST3 TEST4. In addition, the OUTTABLE option outputs the statistics and test results (in the _TESTS_ variable).

```
PROC RAREEVENTS DATA=EVENTS LIMITS=LIMITS_EXTENDED;
      CHART VisitsBetween*Date/
            READPHASES=ALL
```

```
        PHASELEGEND
        LIMITPHASES=ALL
        TEST1 TEST2 (K=8 N=8) TEST3 TEST4
        OUTTABLE=TEST_RESULTS;
RUN;
```

However, the results of this run produce a chart identical to Figure 3 and the TEST_RESULTS table has only null values for the _TESTS_ variable. Figure 3 displays 13 consecutive points above the center line but TEST2 failed to detect it. The reason is that the improvement spans multiple phases. To ignore phase boundaries when applying tests, the TESTACROSS option must be used. In healthcare, TESTACROSS is often appropriate because the impact of a process change might be desired or observed even before a policy becomes official or prior to the date of a technology-related "go-live". This is particularly true with respect to behavior and process changes that are publicized, promoted, or developed within the same group of people who are carrying out an intervention. Adding TESTACROSS to the RAREEVENTS procedure has the effect of signaling a positive TEST2 (Figure 4).

```
PROC RAREEVENTS DATA=EVENTS LIMITS=LIMITS_EXTENDED;
     CHART VisitsBetween*Date/
        READPHASES=ALL
        PHASELEGEND
        LIMITPHASES=ALL
        TEST1 TEST2 (K=8 N=8) TEST3 TEST4
        TESTACROSS
        OUTTABLE=TEST_RESULTS;
RUN;
```
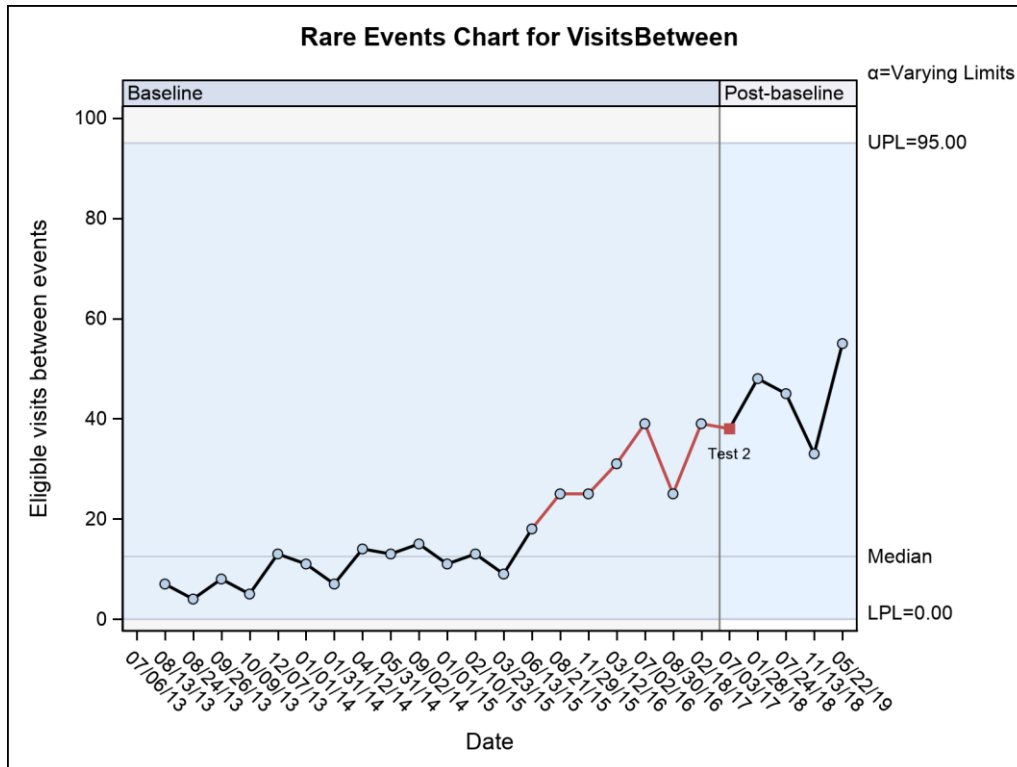

Figure 4. Rare Events Chart with Phases Defined, Baseline Extended, Tests and TESTACROSS Option Applied.

CREATING A SHIFT PHASE

7

These results have provided evidence of a shift. A new phase must be calculated and extended forward into the future so that tests for special cause variation now use the shift period's center line for comparison, rather than that of the baseline. To produce the appropriate rare events chart requires repeating several of the steps above: calculating the center line and limits of the 8-point run, extending it forward from the beginning of those 8 points, defining and naming the shift phase, and obtaining test results prospectively as new occurrences of the event become available. The shift phase should persist for a period of 20 events if available.

The TEST_RESULTS table indicates the eight points that contributed to the TEST2-positive result. The first seven have an asterisk for the _TESTS_ variable and the eighth has _TESTS_ equal to 2. After the eighth point, there are only four more events in the data set. Therefore, our shift period will be 12 points, beginning with the first point in the run. We must replace the values of the _PHASE_ variable in the EVENTS data set with new values.

```
DATA TEST_RESULTS;
      SET TEST_RESULTS;
      REC=_N_;
RUN;

DATA EVENTS;
      SET EVENTS;
      REC=_N_;
RUN;

PROC SQL;
      CREATE TABLE EVENTS AS SELECT
            EVENTS.VisitID, EVENTS.Date, EVENTS.VisitNumber,
EVENTS.VisitsBetween, _TESTS_,
      CASE
            WHEN _TESTS_ IN (' *',' 2')
            OR EVENTS.REC GT (SELECT REC from TEST_RESULTS where _TESTS_ = '
2')
            THEN 'Shift 1 '
      ELSE 'Baseline'
      END
AS _PHASE_
      FROM EVENTS INNER JOIN TEST_RESULTS ON
            EVENTS.REC=TEST_RESULTS.REC;
QUIT;
```

With the new phase defined we are able to produce statistics associated with the baseline and 8-point shift, and output them to the SHIFT_LIMITS data set.

```
PROC RAREEVENTS DATA=EVENTS (WHERE=(_TESTS_ IN (' *',' 2')));
      CHART VisitsBetween*Date/
            DIST=GEOMETRIC
            OUTLIMITS=SHIFT_LIMITS
            READPHASES=ALL
            NOCHART;
RUN;
```

These limits are then combined with the previously calculated statistics for the baseline period, and applied to the rest of the records in the Shift 1 phase via the LIMITS= option.

```
PROC APPEND
      DATA=LIMITS_EXTENDED (WHERE=(_PHASE_='Baseline'))
      BASE=SHIFT_LIMITS
      FORCE;
RUN;
```

8

```
PROC RAREEVENTS DATA=EVENTS LIMITS=SHIFT_LIMITS;
     CHART VisitsBetween*Date/
          READPHASES=ALL
          PHASELEGEND
          LIMITPHASES=ALL
          TEST1 TEST2 (K=8 N=8) TEST3 TEST4
          TESTACROSS
          OUTTABLE=FINAL_RESULTS;
RUN;
```

In the resulting chart, Shift 1 is displayed with its corresponding center line and limits (Figure 5). Prospectively as additional events occur and are added to the ELIGIBLE data set, special cause variation tests will be conducted against the Shift 1 center line.
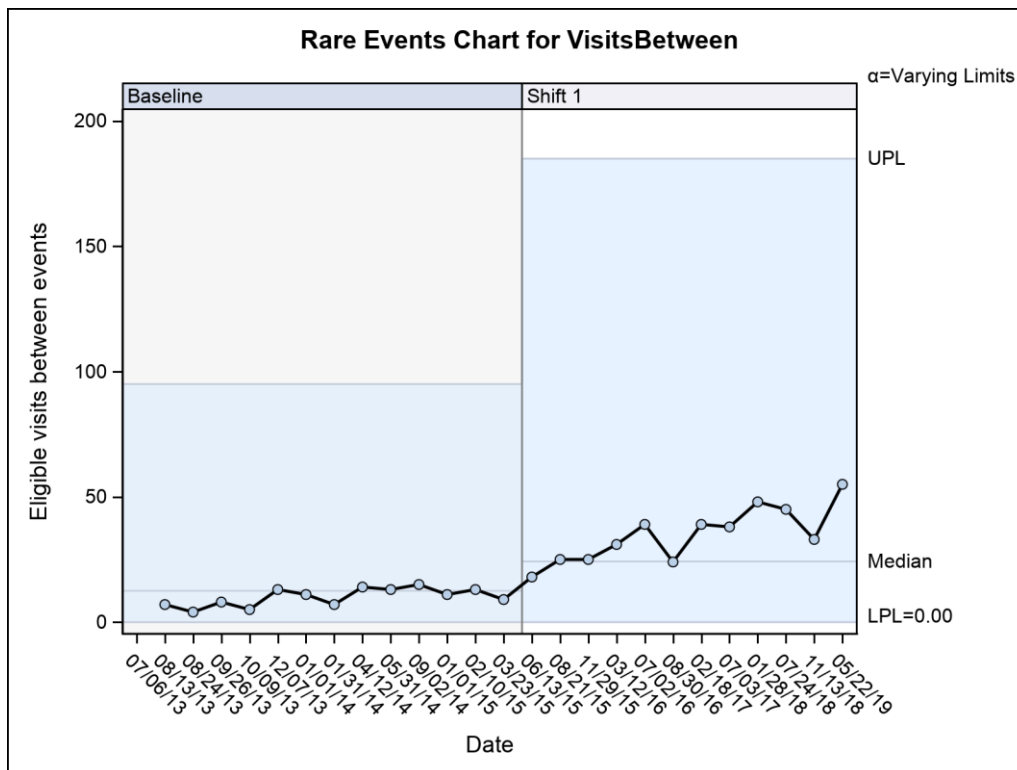
Figure 5. Rare Events Chart with Shift Phase Displayed.

CONCLUSION

Leveraging the options provided by PROC RAREEVENTS facilitates the creation of SPC charts that extend central lines from baseline periods and display shifts to represent process changes. In developing a rare events chart for adverse events, one of the most common scenarios when using SPC in the healthcare setting, these options are essential:

LIMITPHASES, READPHASES=ALL and PHASELEGEND ensure the incorporation of phase definitions and their display in the graphic output.

OUTLIMITS and OUTTABLE generate output datasets containing key statistics that are used in subsequent development steps.

NOCHART suppresses graphic output when the primary purpose of a PROC RAREEVENTS run is to generate output datasets.

TEST1, TEST2, TEST3 and TEST4 distinguish the special cause tests being applied. TEST2 (K=8 N=8) is an example of how the number of points required for a consecutive-point run can be set by the programmer.

TESTACROSS applies special cause tests across phase boundaries.

While our example pertains to a count of opportunities between events, the RAREEVENTS procedure might also be applied to scenarios in which the time between events is measured. Some options, such as DIST=, could vary (Santiago and Smith 2013).

REFERENCES

Benneyan, J.C. 2001. "Design, use, and performance of statistical control charts for clinical process improvement." Accessed August 13, 2019. Available at https://pdfs.semanticscholar.org/87a5/1635f3dca8deddafbfb93e3078e83869596a.pdf.

Benneyan J.C. 2001b. "Performance of Number-Between g-Type Statistical Control Charts for Monitoring Adverse Events." *Health Care Management Science* 4:319–336.

Institute for Healthcare Improvement. "Science of Improvement." Accessed August 13, 2019. Available at http://www.ihi.org/about/Pages/ScienceofImprovement.aspx.

Kaminsky, F. C., Benneyan, J. C., Davis, R. D., and Burke, R. J. 1992. "Statistical Control Charts Based on a Geometric Distribution." *Journal of Quality Technology* 24:63–69.

Provost, L. P. and Murray, S. K. 2011.*The Health Care Data Guide: Learning from Data for Improvement*. San Francisco, CA: John Wiley & Sons.

Ransdell, B. 2016. *Improving Health Care Quality with the RAREEVENTS Procedure,* presented at the SAS Global Forum 2016 Conference in Las Vegas, NV (paper SAS4040-2016).

Santiago, E. and Smith, J. 2013. "Control Charts Based on the Exponential Distribution: Adapting Runs Rules for the t Chart." *Quality Engineering* 25:85–96.

RECOMMENDED READING

- *SAS/QC® 15.1 User's Guide*

- Jones, R. and Liu, L. 2018. *Leveraging SHEWHART Procedure Options to Monitor and Evaluate Improvements in Healthcare,* presented at the Midwest SAS Users Group Meeting in Indianapolis, IN (paper HS-071-2018).

- Liu, L. and Jones, R. 2018. *Use of SAS Macros to Automate the Production of Statistical Process Control Charts,* presented at the Midwest SAS Users Group Meeting in Indianapolis, IN (paper HS-074-2018).

CONTACT INFORMATION

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

> Roderick Jones
> Ann & Robert H. Lurie Children's Hospital of Chicago
> RCJones@luriechildrens.org