Paper 172-2013

# Evaluating System-Wide Process Improvement in a Health-Care System: Data Through Analysis

Eric C. Wong, Lubna S. Qureshi, Pragati Kenkare, Dorothy Hung
Palo Alto Medical Foundation Research Institute

## ABSTRACT

Disruptive system changes are required for sustaining high-quality and affordable health-care delivery systems. Successful, transformative health-care system changes are few and even fewer have been rigorously evaluated. Electronic health records and changes in health IT provide an opportunity to leverage an explosion of data in measuring the impact of process improvement initiatives. This paper provides an example of assessing the impact of a system-wide change in a large, multi-specialty health-care system serving two million patients with a 13-year history of using electronic health records. Lessons from ETL all the way to statistical analysis are detailed including relevant SAS® procedures.

## INTRODUCTION

In light of a recognized, national health-care crisis, disruptive system changes are required for sustaining high-quality and affordable health-care delivery systems. Potential solutions are emerging, but there are few documented, transformative system change solutions, and even fewer have been rigorously studied and evaluated for program effectiveness. Thoughtful program evaluation is necessary for decision makers to identify and scale solutions that are successful, and modify or sunset programs that are not. Electronic health records and improvements in health information technology are providing unprecedented ways to inexpensively and rapidly measure and monitor changes. This paper attempts to comprehensively describe a program evaluation of a system-wide process improvement initiative. We include lessons from defining metrics, data extraction**,** transform, load processes (ETL) through analysis, using one organization's narrative as a case example.

The Lean thinking methodology, a proven process improvement practice in other industries, has been gaining traction in larger healthcare providers. Its core principle is identifying what generates value to the ultimate end user, and scaling more of that while reducing everything else deemed not of value. We describe a program evaluation of one large, multi-center, group practice healthcare organization that began piloting Lean thinking in primary care at one of its locations in November 2011. Two comparable locations were selected as contemporaneous controls. After defining metrics, we extracted data from January 2011- December 2012, and analyzed the time series.

## METRICS

Defining appropriate metrics is an art and continuous process improvement evaluation. For this evaluation, with input from operation staff and researchers, metrics for patient access, organizational affordability, organizational cost, physician production, operational defects, quality of care, and satisfaction were all defined. Our greatest lesson around defining metrics is to document and build consensus among physicians, operations staff, data stewards, researchers, and statistical analysts. In this paper, we present two metrics as examples.

### Metric 1: Active Patient List (APL)

Each primary care physician (PCP) has a number of patients they are serving, called the panel size, and is best approximated using the active patient list (APL). The APL is defined as the number of patients on a PCP's panel with billing activity in primary care in the last 24 months, adjusted by the amount of time the physician worked as measured by clinical full-time equivalent (cFTE). If the process improvement is effective and APL is increasing over time, we could conclude the PCP is taking care of more patients in the same amount of time. Alternatively, if the process improvement is effective and APL is decreasing over time, we might conclude the PCP is taking care of only those patients who need the level of care only the PCP can provide.

### Metric 2: Relative value unit (RVU) production

In a fee-for-service setting, services rendered by physicians are assigned fees. To compare services across physician, department, or organization, it is necessary to have a standard fee schedule. We chose to use Medicare's relative value units. RVUs are a measure of value used in the Medicare reimbursement formula for physician services. The services are classified under a nomenclature based on the Current Procedural Terminology (CPT codes). Each service in the fee schedule is scored under the resource-based relative value scale (RBRVS) to

determine a payment. There are three types of RVUs: physician work RVU (wRVU), malpractice RVU (mRVU), practice expense RVU (pRVU), and total RVU which is the summation of wRVU, mRVU, and pRVU. Each of the three RVUs for a given service (CPT code) is multiplied by a unique geographic practice cost index. If the process improvement is effective and we observe increases in wRVUs and tRVUs over time, we could conclude physicians are rendering more services in the same amount of time.

### STATISTICS

The main analytic method employed was segmented regression for an interrupted time series. An interrupted time series is a time series that has been interrupted, in our case, by a process improvement intervention, and can be divided into two or more segments. The first segment includes observations prior to the intervention, which we call the pre-segment, and help to establish a baseline trend. The intervention occurs, then the second segment includes observations following the start of the intervention, which recalled the post segment, and establishes a post-intervention trend. See Carroll 2008, Wagner et al 2002, and Lagarde 2012 for step-by-step details on performing such an analysis.

## DATA EXTRACTION AND PREPARATION

As likely with most healthcare providers, the data for analysis was distributed across several disparate systems primarily used for scheduling, billing transactions, and clinical encounters. Electronic health records, while mostly synonymous with systems to manage clinical encounters, may also include systems for scheduling and billing. We provide several key lessons from our process of data extraction, transformation, and loading (ETL), validation, and preparation of analytic data sets.

#### Lessons during ETL

1. **Create unified dimension tables**. When trying to harmonize data from disparate relational databases, it was integral to have unified dimension tables. We developed dimension tables for patients, providers, and procedures that were the distinct union of sets from different source systems. While laborious, this may be one time cost.
2. **Validate table keys**. When using data from different sources, it was integral to ensure unique identifiers and primary keys are uniform throughout. For example, in order to identify whether or not a physician in one source system is correctly identified using the same identifier in another system, it was important for both dimension tables to have as much identifying information for validation as possible. This could be national provider identifier (NPI) number for linking providers, social security number for linking patients, or even names. Whenever possible, we recommend using a numeric key, instead of a character variable as a key, such as an individual's name.
3. **Leverage and understand existing data**. To generate results that have the greatest value, it was important for us to understand existing data marts, queries, reports, and analyses previously completed. In disseminating results, it was also important to contextualize our results with what had historically been presented.
4. **Validate early and often.** For every step during ETL, we validated the incremental construction of the data by comparing attributes to our expectations. This included observation or record counts, examining data volume (clinical encounters or transactions) over time, examining top 10 subsets, and univariate simple statistics. (PROC COMPARE, PROC MEANS, PROC UNIVARIATE). When merging or joining tables or data sets, it was integral to compare variable attributes prior to the join to ensure that variables are not truncated in the process.

#### Lessons when preparing analytic data sets

1. **Standardize data to minimize confounding.** Often, there may be system changes, definition changes, and other confounders that may affect data over time. For example, the Centers for Medicare & Medicaid Services (CMS) is responsible for maintaining the fee schedule and makes periodic changes to the assignment of RVUs to procedures. This may mean in one year, a procedure generates 1.3 RVUs, but in the next year, the same procedure may generate to 2.1 RVUs. For all analyses with RVUs, we restated to the CMS's 2012 second release of RVUs valuation.
2. **Work with the lowest grain, then aggregate.** We found it most helpful to work with data in its lowest grain first, and then aggregate to the level of analysis at the end. For example, if the main comparison was across departments, but the data is available at the physician, or encounter level, we found it most useful to transform and validate the data at its lowest level, then aggregate up to the unit of analysis.
3. **Understand what data is included.** When preparing analytic data sets, exclusion criteria are often explicit; however, it is also important to understand what data is being implicitly included. For example, we gathered physician level data from specific locations and departments from a cross-section of time. During this time, physicians were hired and also left the organization and this may have accounted for confounding in our metrics. Therefore, we added an additional criterion of continuous employment during the time.

## VISUALIZATION

First, we generated plots of our metric over time using the SAS/GRAPH® SG Procedures. Below, each panel represents a primary care department (Family Medicine, Internal Medicine, Pediatrics). Each of the series represents the monthly values of a clinic location. The intervention began in Location 1 at November 2011 as indicated by the red vertical line. Location 2 and 3 are used as contemporaneous controls and did not have the process improvement intervention.

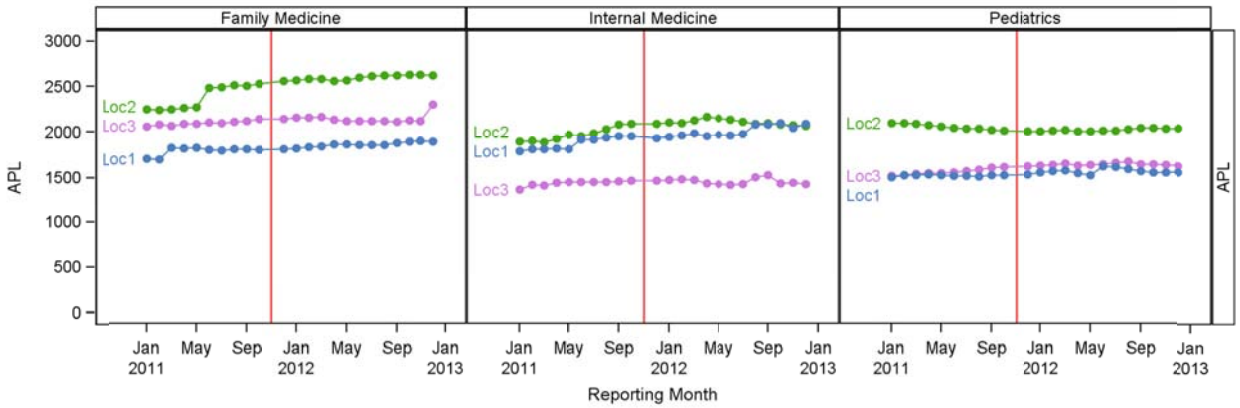


**Figure 1: Plot of active patient list (APL) by location and department**

```
ODS HTML FILE='apl_sas.html' PATH='.' STYLE=journal IMAGE_DPI=200;
ODS GRAPHICS / RESET WIDTH=9in HEIGHT=3in NOBORDER;

PROC SGPANEL DATA=work.fig1_apl;
       PANELBY department metric / LAYOUT=LATTICE ONEPANEL COLUMNS=3 NOVARNAME;
       REFLINE '01NOV2011'D /
              AXIS=X
              LINEATTRS=(THICKNESS=0.25px PATTERN=SOLID COLOR=RED);
       SERIES
              X=date_report_mmyy
              Y=loc3 /
              MARKERS
              MARKERATTRS=(SIZE=3PX SYMBOL=CIRCLEFILLED COLOR=VLIP)
              LINEATTRS=(COLOR=VLIP PATTERN=SOLID THICKNESS=1pt)
              CURVELABEL="Loc3"
              CURVELABELPOS=MIN;
       SERIES
              X=date_report_mmyy
              Y=loc2 /
              MARKERS
              MARKERATTRS=(SIZE=3PX SYMBOL=CIRCLEFILLED COLOR=VILG)
              LINEATTRS=(COLOR=VILG PATTERN=SOLID THICKNESS=1pt)
              CURVELABEL="Loc2"
              CURVELABELPOS=MIN;
       SERIES
              X=date_report_mmyy
              Y=loc1 /
              MARKERS
              MARKERATTRS=(SIZE=3PX SYMBOL=CIRCLEFILLED COLOR=BIGB)
              LINEATTRS=(COLOR=BIGB PATTERN=SOLID THICKNESS=1pt)
              CURVELABEL="Loc1"
              CURVELABELPOS=MIN;

       ROWAXIS MIN=0 MAX=3000 LABEL="APL";
       COLAXIS MIN='01JAN2011'D LABEL="Reporting Month";
RUN;
```

## BUILDING THE SEGMENTED REGRESSION MODEL

First, we performed a segmented regression on the location that received the intervention, Location 1. Later, we extend the approach to incorporate contemporaneous controls.

We investigated the autocorrelation by examining plots of the partial autocorrelation function (PACF). One way to produce these plots is to use the SAS/ETS® procedure PROC ARIMA with the IDENTIFY statement.

```
PROC ARIMA DATA=...;
       BY metric dept_ordered;
       IDENTIFY VAR=loc1 STATIONARITY=(DICKEY=0);
RUN;
```

We also investigated seasonality by looking at the spectral density plots using PROC SPECTRA. Since PROC SPECTRA does not produce any output by default, we used the SG procedures to create the plots.

```
PROC SPECTRA DATA=... =... CENTER S;
       BY metric department;
       VAR loc1;
       WEIGHTS PARZEN;
RUN;
PROC SGPANEL DATA=...;
       PANELBY department metric / LAYOUT=LATTICE ONEPANEL COLUMNS=4 NOVARNAME;
       SERIES
              X=period
              Y=s_01 /
              MARKERS
              MARKERATTRS=(SIZE=3PX SYMBOL=CIRCLEFILLED COLOR=BLACK)
              LINEATTRS=(COLOR=BLACK PATTERN=SOLID THICKNESS=1pt)
              CURVELABEL="LOC1"
              CURVELABELPOS=START;
       ROWAXIS LABEL="Spectral Density";
       COLAXIS MIN=1 MAX=24 VALUES=(1, 3, 6, 9, 12, 15, 18, 21, 24) LABEL="Period";
RUN;
```

In our example with APL, there is a first-order autocorrelation structure, but no seasonality detected. We built the model using PROC AUTOREG with the NLAG option on the MODEL statement. The specification used is the same as described in Wagner et al 2002 and Carroll 2008.

```
PROC AUTOREG DATA=...;
       BY metric department;
       MODEL loc1 = time intervention time_aft_int / NLAG=(1);
RUN;
```

## INCORPORATING CONTEMPORANEOUS CONTROLS

Contemporaneous controls, when available, are important for interpreting results of the segmented regression. For example, one might find in the intervention location that there is a statistically significant difference in the slope of the series after the intervention began. If however the contemporaneous control exhibits the same pattern, then there is less confidence that the difference in the post period can be attributable to the intervention. In contrast, if the contemporaneous control exhibits no change in the post period, as expected from a control, then there is greater confidence in concluding there is a difference between intervention and control locations following the intervention.

While there are many approaches to measuring the difference between the intervention and contemporaneous control series, we decided to measure this by examining the differenced series between the intervention location and each of the contemporaneous controls. That is, for each month, we generated a series by subtracting the value from the intervention location and the control location (Loc1 – Loc2) or (Loc1 – Loc3), respectively. Then, we applied the same methodology for segmented regression as described above, but on each of the two differenced series.

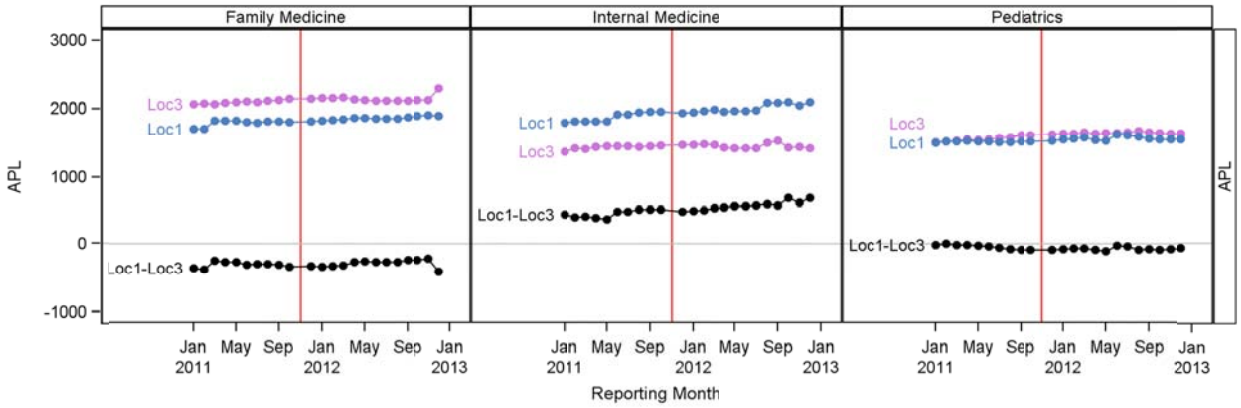


**Figure 2: Plot of active patient list (APL) by department in two locations (Loc1, Loc3), and the differenced series.**

The differenced series as represented above in black is the point-wise differences between the two series. We may use the same code as above when analyzing just Location 1, and the same specifications for variables related to time, and simply replace the variable `loc1` with a new differenced variable `loc1_loc3` whenever applicable.

```
PROC AUTOREG DATA=...;
      BY metric department;
      MODEL loc1_loc3 = time intervention time_aft_int / NLAG=(1);
RUN;
```

Statistically significant model parameter estimates, especially the beta coefficients for `intervention` and `time_aft_int`, indicate whether the two series are different from each other. The coefficient for intervention indicates the intercept for the post segment, meaning whether the two series are different in post segment immediately following the intervention. And the coefficient for `time_aft_int` indicates the slope for the post segment, meaning whether the two series have different slopes in the post..

## ANOTHER EXAMPLE: RVUS

We provide another example of the segmented regression for interrupted time series by examining our second metric RVUs; in particular we will focus on total RVUs (tRVUs). For each department in each location, we calculated tRVUs per clinical full time equivalent (cFTE), to account for physicians that are working both full-time and part-time.

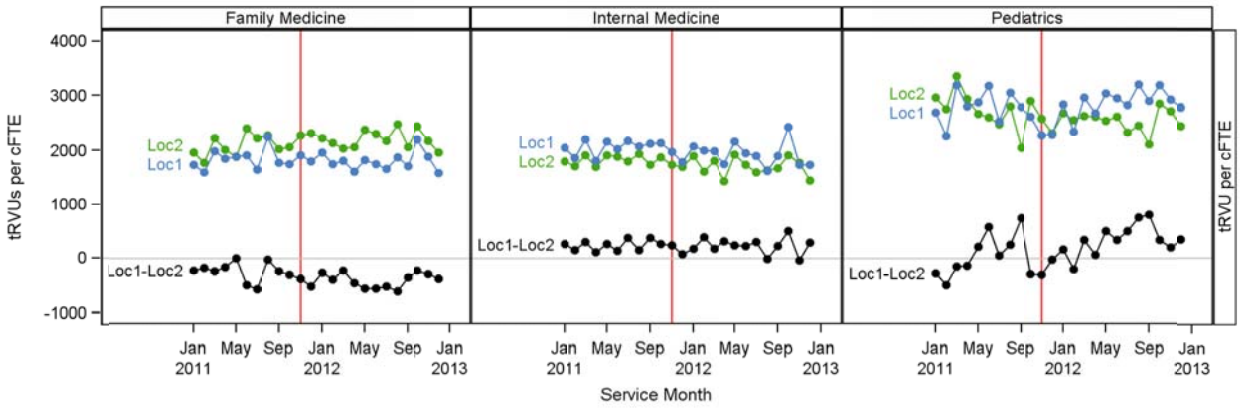


**Figure 3: Plot of tRVUs by department in two locations (Loc1, Loc2), and the differenced series.**

In this example, we built separate models for the departments as some had evidence of a first-order autocorrelation structure, while others did not. The approach otherwise was identical as before.

## CONCLUSION

When conducting quantitative program evaluation, careful thought has to be put in every step from the definition of metrics, the data extraction transformation and loading process, preparation of analytic data sets, and analytic methodology. This paper provided the narrative on two metrics that were part of a comprehensive quantitative evaluation of a system-wide process improvement intervention.

## REFERENCES


- Carroll, Nikki. 2008. "Application of Segmented Regression Analysis to the Kaiser Permanente Colorado Critical Drug Interaction Program." *Proceedings of the Western Users of SAS® Software 2008 Conference*. Universal City, CA: SAS. Available at: http://www.wuss.org/proceedings08/08WUSS%20Proceedings/papers/anl/anl08.pdf
- Wagner, AK, Soumerai, SB, Zhang, F, and Ross-Degnan, D. 2002. "Segmented regression analysis of interrupted time series studies in medication use research." *Journal of Clinical Pharmacy and Therapeutics*. 27, 299-309.
- Lagarde, Mylene. 2012. "How to do (or not to do) … Assessing the impact of a policy change with routine longitudinal data." *Health Policy and Planning*. 27 (1): 76-83.


## RECOMMENDED READING


- *SAS/ETS® User's Guide*
- Chvosta, Jan and Little, Mark. 2009. "Model Fitting and Data Analyses in SAS/ETS® Software Using ODS Statistical Graphics." *Proceedings of the SAS® Global Forum 2009 Conference.* Washington, DC: SAS. Available at: http://support.sas.com/resources/papers/proceedings09/243-2009.pdf
- Toussaint, John and Gerard, Roger. 2010. *On the Mend: Revolutionizing Healthcare to Save Lives and Transform the Industry.* Cambridge, MA: Lean Enterprises Institute.


## CONTACT INFORMATION

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


Eric C. Wong
Palo Alto Medical Foundation Research Institute
795 El Camino Real
Palo Alto, CA 94301
wonge@pamfri.org