Application of Segmented Regression Analysis to the Kaiser Permanente Colorado Critical Drug Interaction Program

Nikki Carroll, Kaiser Permanente, Denver, CO

ABSTRACT

Failure to detect significant drug interactions may result in adverse outcomes. While proper screening and management of drug interactions can prevent the majority of adverse events, studies indicate that current practice is suboptimal. In the last quarter of 2001, physicians and pharmacists in Kaiser Permanente Colorado (KPCO) developed an electronic critical drug interaction alert program. Electronic screening was coupled with active intervention to prevent dispensing of critically interacting drug combinations.

Rates of critical drug interactions were collected 20 months pre-intervention and 37 months post-intervention. Success of the intervention was based on the changes in the rates of critical drug interactions and statistically evaluated by segmented regression analysis (a powerful statistical method for estimating intervention effects in interrupted time series).

This paper shows the application of segmented regression analysis to the critical drug interaction program.

KPCO CRITICAL DRUG INTERACTION PROGRAM BACKGROUND

A drug-drug interaction occurs when the effects and/or kinetics of one drug are altered by the co-administration of a second drug. The incidence of clinically significant interactions in the outpatient setting ranges from 0.6% to 23.3%. Failure to detect significant drug-drug interactions may result in adverse outcomes for patients and increased health care costs. It has been reported that drug interactions cause up to 2.8% of hospitalizations each year, resulting in an estimated annual cost of $1.3 billion.

Although proper screening and management of drug-drug interactions can prevent the majority of adverse drug events, studies indicate that current practice is sub-optimal. A large volume of clinically insignificant alerts can result in “alert fatigue”, and clinically significant alerts may be overridden. In the last quarter of 2001, physicians and pharmacists in Kaiser Permanente Colorado (KPCO) collaborated to develop a program to identify and address critically interacting drugs. To minimize the risk of bypassing significant alerts, electronic screening was coupled with active intervention to prevent dispensing of drug combinations that have the potential for serious adverse interactions.

The objective of this study was to assess the impact of the Critical Drug Interaction program (CDIX). We hypothesized that compared to a baseline period the rate of co-dispensing of drugs that critically interact in our health care system would be lower after the introduction of CDIX. Eight drug combination pairs were chosen for analysis. These drug pairs were combinations that both pharmacists and physicians agreed were “never use” drug-drug combinations and had appropriate alternatives available.

Monthly electronic outpatient pharmacy data were collected on these eight drug pairs between July 2000 and May 2005; this time period includes 20 months before the program was implemented and 37 months after it was implemented. KPCO members who received an outpatient dispensing from the medical office pharmacies for any of the eight drug pairs of interest during the study period were included in the analysis. Segmented regression analysis was used to estimate changes in the rates of critical drug interactions. Analyses were conducted using SAS® 9.1 (SAS Institute, Carey NC).

TIME SERIES REVIEW

Time series refers to a large series of observations made on the same variable consecutively over time. Time series analyses have traditionally been used for forecasting techniques in economics and business. For example, forecasting techniques are used to predict product sales over time, employment rates over time or product inventory over time. However, time series analysis may also be used to evaluate therapeutic effects in health care.

INTERRUPTED TIME SERIES

Interrupted time series (ITS) is a special kind of time series that can be used to measure a treatment effect or the impact of an intervention. The goal is to demonstrate a clear causal relationship between an intervention and an outcome after ruling out other forces that might have had the same outcome in the absence of the intervention.

ITS are divided into at least 2 segments separated by an intervention. The first segment includes a series of pre-intervention observations that establish a baseline trend. The intervention occurs at a known time and then is followed with a series of post-intervention observations from which we can analyze the impact of the intervention.

Generally, 12 data points before the intervention and 12 data points after the intervention are needed which allow for the detection of seasonal variation within the data. In addition, it is recommended to collect approximately 1,000 data points at each data point in order to achieve an acceptable level of variability.

In our study, we were looking at the impact of the CDIX intervention on the rate of critically interacting drugs. We collected 20 months of data prior to the intervention and 37 months of follow up data after the intervention in order to detect seasonality.

SEGMENTED REGRESSION ANALYSIS

Segmented regression analysis is a powerful method for estimating how much an intervention affects the outcome measure immediately and over time. Segmented regression models fit a least squares regression line in each segment and assumes a linear relationship between the independent variable and the outcome within each segment. Data must be collected at equally spaced intervals over time for a segmented regression analysis. Model 1 shows how a segmented regression will estimate changes in the rates of co-prescribing before the intervention and changes in the rates of co-prescribing after the intervention in our study:

$$Rate_t = \beta_0 + \beta_1 * \text{time}_t + \beta_2 * \text{intervention}_t + \beta_3 * \text{time after intervention}_t + e_t \quad (1)$$

Where:

- $Rate_t$ is the rate of co-prescribing of critically interacting drugs
- β_0 estimates the baseline co-prescribing rate at the beginning of the study period
- β_1 estimates the change in co-prescribing rates that occur with each month before the intervention
- *time* is a continuous variable indicating the number of months prior to and after the intervention. It ranges from -20 months to 37 months
- β_2 estimates the change in the co-prescribing rate immediately following the intervention
- *intervention* indicates whether or not the intervention had taken place during that time period (before the intervention is coded as $\text{intervention}_t = 0$ and after the intervention is coded as $\text{intervention}_t = 1$)
- β_3 estimates the change in the slope after the intervention compared to the slope before the intervention
- *time after intervention* is a continuous variable indicating the number of months that have passed since the intervention was implemented. This is coded as zero for all time periods prior to the intervention
- e_t represents the random error.

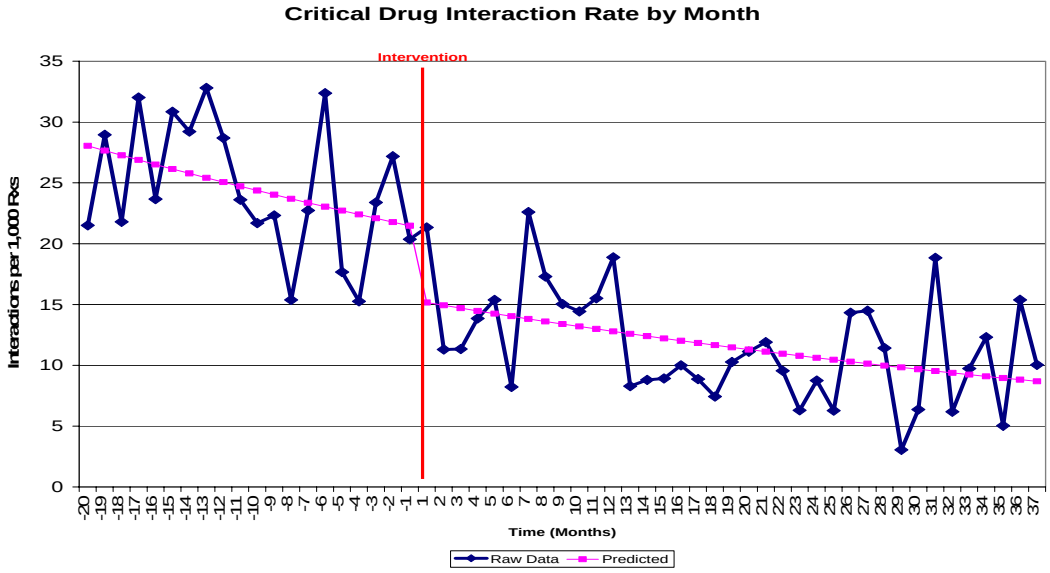
There are several components to consider when building our segmented regression model: visual inspection of the data, lagged effects, autocorrelation, and seasonality.

BUILDING THE MODEL

VISUAL INSPECTION OF DATA

The graphical representation of the intervention is the first step when analyzing interrupted time series data. This allows you to determine whether an effect is immediate or delayed, abrupt or gradual, and whether or not an effect is sustained over time. The rates of co-prescribed critically interacting drugs are compared pre-intervention with post-intervention, looking for changes in the series patterns. Figure 1 shows the CDIX intervention program rate by month.

Figure 1. Critical Drug Interaction Rate by Month



Also graphed is a "predicted regression line" to better see the trend of co-prescribing rates. The predicted line was not analyzed; it was calculated for the visual representation only. The predicted line was calculated as follows:

```
proc genmod data=crit_intx;
  model num_intx/totrx = intervention intervention*time/
        link=logit dist=binomial noint;
  output pred=pred;
run;
```

Our graph shows a decreasing trend in the slope before the intervention, an abrupt drop in the rate immediately following the intervention, and then a gradually decreasing slope continues after the intervention. These will be tested statistically once our final model is built.

LAGGED EFFECTS

The effect of an intervention may take time to appear. The effect may occur several time periods after the intervention. It is important to account for this "lag" in the analysis in order to avoid incorrect specification of the intervention effects. Lag periods can be excluded from the analysis or they can be analyzed as a separate segment in the model.

In our study the intervention was immediate through the electronic pharmacy system, therefore, no lag effects existed and they were not entered into our model.

AUTOCORRELATION

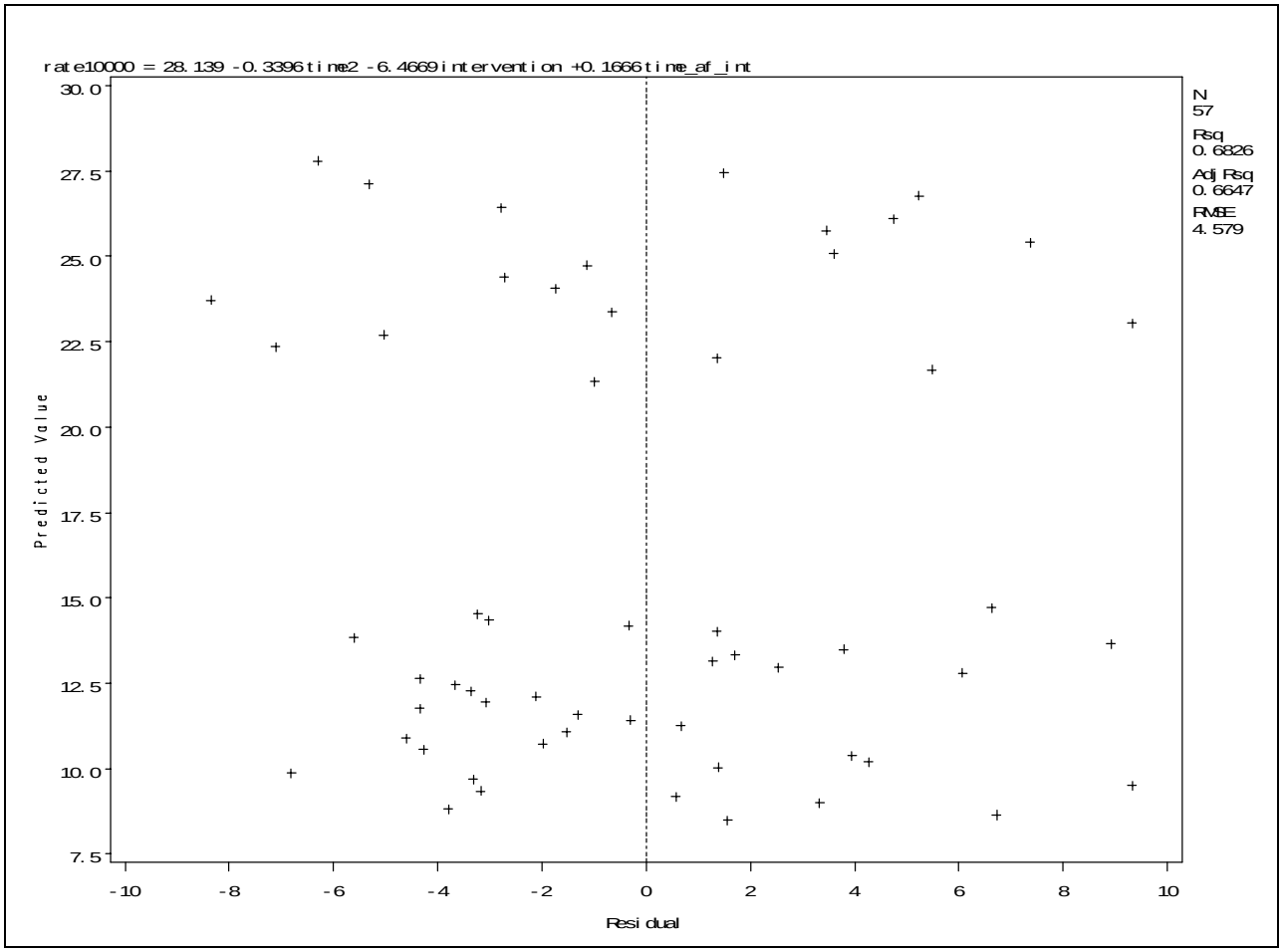
Ordinary least squares regression analysis assumes that error terms associated with each observation are uncorrelated. Prescribing patterns and other health outcomes at two time points that are close to each other may be more similar than outcomes at two time points further apart, resulting in serial autocorrelation of the error terms. Correlation between adjacent data points is termed first-order correlation; correlation between the current point and two months before or after would be second-order autocorrelation and so forth. Failing to correct for autocorrelation may lead to underestimated standard errors and overestimated significance of the effects of an intervention. If autocorrelation exists, we can control for it in our model.

Autocorrelation can be detected visually by inspecting a plot of the residuals over time and by conducting statistical tests. Randomly scattered residuals with no discernable pattern indicate no autocorrelation. Positive correlation exists when consecutive residuals lie on the same side of the regression line; negative autocorrelation exists when consecutive residuals tend to lie on the different sides of the regression line.

The residuals were graphed using the following code:

```
proc reg data=crit_intx;
  model rate = time intervention time_af_int/dw;
  plot rstudent.*obs.
    /vref= -1.714 1.714 cvref=blue  lvref=1
       href = 0 to 60 by 5 chref=red cframe=ligr;
  plot predicted.*residual.;
run;
```

which produced the residual plot below. The plot shows no discernable pattern in the residuals:



Autocorrelation can also be detected by the Durbin-Watson statistic. The Durbin-Watson statistic tests for serial correlation of the error terms in the regression model. Values close to 2.0 indicate no serious autocorrelation. If the statistic is significant, the model can be adjusted by estimating the autocorrelation parameter and including it in the segmented regression model.

PROC AUTOREG has an option to use the Durbin-Watson statistic and test for autocorrelation:

```
proc autoreg data=crit_intx;
  model rate=time intervention time_af_int/ dwprob;
run;
```

The Durbin-Watson statistic for the regression model of the CDIX was 2.0173 (p-value for hypothesis of negative autocorrelation = 0.6356, p-value for hypothesis of positive autocorrelation = 0.3644), indicating no autocorrelation and confirming our visual inspection of the residual graph.

Ordinary Least Squares Estimates

SSE	1111.28298	DFE	53
MSE	20.96760	Root MSE	4.57904
SBC	347.233692	AIC	339.061487
Regress R-Square	0.6826	Total R-Square	0.6826
Durbin-Watson	2.0173	Pr < DW	0.3644
Pr > DW	0.6356		

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

SEASONALITY

Time series sometimes exhibit seasonal fluctuations. For example, prescribing in January of one year is more similar to prescribing in January a year ago than to prescribing in other months. Another example would be variation in drug utilization due to treatment of illnesses that vary by season, therefore, in order to detect seasonality, at least 24 monthly data points are required. If seasonality exists, it is important to control for it when estimating intervention effects so that the intervention effects are more likely to represent true intervention effects.

Another way to look at seasonality is to look at whether or not our data is stationary. If a series is stationary then the mean is constant over time, the variance of our outcome is constant over time, and the covariance between our outcome at different time periods must match. If the series has a seasonality or some other non-stationary pattern, the usual solution is to take the difference of the series from one period to the next and then analyze this differenced series. Sometimes a series may need to be differenced more than once or differenced at lags greater than one period.

We tested for seasonality/stationarity using the Dickey-Fuller unit root test in PROC ARIMA using the following code:

```
proc arima data=crit_intx;
  identify var=rate stationarity=(dickey=0);
quit;run;
```

The identify statement specifies the time series to be modeled and "var" names the variable containing the outcome to analyze. The stationarity option with a dickey test is specified next. A value of zero lets SAS automatically compute the lags that would be necessary.

The null hypothesis is tau is not stationary. The output shows the following results. In looking at the "Single Mean" line, we conclude that the series is stationary and we do not need to correct for seasonality/stationarity with differencing.

Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3.9062	0.1701	-1.53	0.1179		
Single Mean	0	-17.5830	0.0137	-3.16	0.0274	5.04	0.0404
Trend	0	-46.0607	0.0001	-6.10	<.0001	18.62	0.0010

ANALYSIS AND INTERPRETATION OF THE MODEL

We learned in the previous section that we did not need to correct for lagged effects, autocorrelation or seasonality. Therefore, our final model will be specified exactly how we defined it for Model 1 above:

$$Rate_t = \beta_0 + \beta_1 * \text{time}_t + \beta_2 * \text{intervention}_t + \beta_3 * \text{time after intervention}_t + e_t \quad (1)$$

We will interpret the "level" and "trend" in our results and along with the visual graph make conclusions.

LEVEL

The level is the value of the series at the beginning of a given time interval (i.e., on our graph, it is the intercept of the series on the y axis). If there is a change in the values following intervention, the change is called a change in level or intercept, because 1) the level of the series drops and, 2) the pre- and post-treatment slopes would have different intercepts. This would constitute an intervention effect.

TREND

The trend is the rate of change of a measure or the change in slope. A change in trend is observed by an increase or decrease in the slope of the segment after the intervention as compared with the segment preceding the intervention.

Although differences in trend and level can be detected visually, statistical tests still need to be run in order to detected if the differences are the result of chance alone or factors other than the intervention.

PROC AUTOREG MODEL

Our model was specified in PROC AUTOREG to estimate the changes associated with the intervention while controlling for baseline values.

```
proc autoreg data=crit_intx outest=est;
  model rate=time intervention time_af_int/method=ml;
run;
```

The output from the model reveals:

Final ITS Model

The AUTOREG Procedure

Dependent Variable rate

Ordinary Least Squares Estimates

SSE	1111.28298	DFE	53
MSE	20.96760	Root MSE	4.57904
SBC	347.233692	AIC	339.061487
Regress R-Square	0.6826	Total R-Square	0.6826
Durbin-Watson	2.0173		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	28.1392	2.1271	13.23	<.0001
time	1	-0.3396	0.1776	-1.91	0.0612
intervention	1	-6.4669	2.5010	-2.59	0.0125
time_af_int	1	0.1666	0.1911	0.87	0.3872

OUTPUT

The *intercept* variable (measuring our level) shows that just before the beginning of the observation period, the rate of critical drug interactions was 28.1 per 10,000 prescriptions. The *time* variable (measuring our trend) shows that before the intervention, there was no significant month-to-month change in our trend, i.e., the mean number of critical drug interactions (p-value for baseline trend = 0.0612). The *intervention* variable (measuring our level after the intervention), shows that immediately following the intervention the rate of critical drug interactions significantly dropped by 6.5 prescriptions per 10,000 (p=0.0125). The *time_af_int* variable (measuring our trend after the intervention – sustainability) shows no significant change in the month-to-month trend in the mean number of critical drug interactions after the intervention (p-value for trend change = 0.3872). After stepwise elimination of non-

significant terms, the most parsimonious model contained only the intercept and the significant level change in the mean number of critical drug interactions.

ADVANTAGES

Although the gold standard for study design is the randomized control trial, it's not always feasible or ethical to randomize patients or find suitable controls. Interrupted time series is a great alternative since patients can serve as their own controls and, in general, you have the power to test and correct for seasonal patterns and outliers. The visual representation of interrupted time series allows you to see the response to an intervention and whether the effect of that intervention can be sustained over time. Segmented regression analysis easily allows you to control for prior trends and analyze the response to an intervention. Estimating the size of the effect at different time points and changes in the trend over time can also be estimated using segmented regression analysis.

DISADVANTAGES

The model used in this paper assumed the outcome in each segment followed a linear trend. Linear trends may not hold over long time periods and changes may follow non-linear patterns. Although it is recommended that approximately 1,000 data points be collected in each time period, our outcome was rare and we did not collect that many.

CONCLUSION

Segmented regression analysis of interrupted time series data is a robust modeling technique that allows the analyst to estimate dynamic changes in various processes and outcome following interventions intended to change medication use, while controlling for secular changes that may have occurred in the absence of the intervention.

The use of the segmented regression analysis showed that employing an intervention system that limits electronic alerts regarding drug-drug interactions to those deemed critical but also required pharmacist intervention and collaboration with the prescriber decreases the number of critical drug interaction dispensed.

REFERENCES

Bergendal L, Friberg A, Schaffrath A. Potential drug-drug interactions in 5,125 mostly elderly out-patients in Gothenburg Sweden. *Pharm World Sci* 1995;17:152-7.

Bjerrum L, Andersen M, Petersen G, Kragstrup J. Exposure to potential drug interactions in primary health care. *Scand J Prim Health Care* 2003;21:153-8.

Bjorkman IK, Fastbom J, Schmidt IK, Bernsten CB, the Pharmaceutical Care of the Elderly in Europe Research (PEER) Group. Drug-drug interactions in the elderly. *Ann Pharmcother* 2002;36:1675-81.

Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. *JAMA* 1997;277:301-6.

Costa AJ. Potential drug interactions in an ambulatory geriatric population. *Fam Pract* 1991;8:234-6.

Dickey, David A. Stationary Issues in Time Series Models. *SUGI 30 Proceedings, paper 192-30*.

Dickey, David A. Case Studies in Time Series. *SUGI28 Proceedings, paper 252-28*.

England E. How Interrupted Time Series Analysis Can Evaluate Guideline Implementation. *The pharmaceutical Journal.* 2005;275:344-347.

Grissinger M. Optimizing the use of computer system clinical alerts. *P&T* 2002;27:593.

Hamilton RA, Briceland LL, Andritz MH. Frequency of hospitalization after exposure to known drug-drug interactions in a Medicaid population. *Pharmacotherapy* 1998;18:1112-20.

Hartshorn EA, Tatro DS. Principles of drug interactions. In: Tatro DS, ed. Drug interaction facts. 5th ed. St. Louis: Facts and Comparisons;1996:xix.

Humphries TL, Carroll NM, Chester EA, Magid DJ, Rocho B. Evaluation of an Electronic Critical Drug Interaction Program Coupled with Active Pharmacist Intervention. *Ann Pharmacother*. 2007 Dec;41(12):1979-85.

Jankel CA, Speedie SM. Detecting drug interactions: a review of the literature. *Ann Pharmacother* 1990;24:982-9.

Jankel CA, Fitterman LK. Epidemiology of drug-drug interactions as cause of hospital admissions. *Drug Saf* 1993;9:51-9.

Johnson JA, Bootman JL. Drug-related morbidity and mortality. A cost-of-illness model. *Arch Intern Med* 1995;155:1949-56.

Juurlink DN, Mamdani M, Kopp A, Laupacis A, Redelmeier DA. Drug-drug interactions among elderly patients hospitalized for drug toxicity. *JAMA* 2003;289:1652-8.

Matowe LK, Leister CA, Crivera C, Korth-Bradley JM. Interrupted Time Series Analysis in Clinical Research. *The Annals of Pharmacotherapy*. 2003;37(7):1110-1116.

Monane M, Matthias DM, Nagle BA, Miriam A. Improving prescribing patterns for the elderly through an online drug utilization review intervention: a system linking the physician, pharmacist, and computer. *JAMA* 1998;280:1249-52.

Peng CC, Glassman PA, Marks IR, Fowler C, Catiglione B, Good CB. Retrospective drug utilization review: incidence of clinically relevant potential drug-drug interactions in a large ambulatory population. *J Managed Care Pharm* 2003;9:513-22.

Roblin DW, Leape LL. Pharmacies and prevention of potentially fatal drug interactions. *JAMA* 1996;275:1086-7.

Rosholm JU, Bjerrum L, Hallas J, Worm J, Gram LF. Polypharmacy and the risk of drug-drug interactions among Danish elderly. *Dan Med Bull* 1998;45:210-3.

SAS Institute Inc. 2006. SAS OnlineDoc® 9.1.3. Cary, NC: SAS Institute Inc.

Schneitman-McIntire O, Farnen TA, Gordon N, Chan J, Toy WA. Medication misadventures resulting in emergency department visits at an HMO medical center. *Am J Health-Syst Pharm* 1996;531:416-22.

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented Regression Analysis of Interrupted Time Series Studies in Medication Use Research. *J Clinical Pharmacy and Therapeutics.* 2002;27:299-309

Zhang F, Wagner A, Soumerai SB, Ross-Degnan D. Estimating Confidence Intervals Around Relative Changes in Outcomes in Segmented Regression Analyses of Time Series Data. *Proceedings of the Fifteenth Annual North East SAS User Group Conference 2002*.

ACKNOWLEDGMENTS

I relied heavily on Anita Wagner et al and Zhang et al for their papers on segmented regression analysis. I have used their models for my analysis here and if not for their work and publications, I would not have been able to so easily interpret my results.

CONTACT INFORMATION

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

Nikki Carroll
Kaiser Permanente
P.O. Box 378066
Denver, CO 80237
Work Phone: (303) 614-1251
E-mail: nikki.m.carroll@kp.org