PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2015; **24**: 663–668
Published online 22 April 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.3783

ORIGINAL REPORT

Signal detection using change point analysis in postmarket surveillance†

Zhiheng Xu[1]*, Taha Kass-Hout[2], Colin Anderson-Smits[3] and Gerry Gray[1]

[1]*Division of Biostatistics, U.S. Food and Drug Administration, Silver Spring, MD, USA*
[2]*Chief Health Informatics Officer, Chief Technology Officer, Office of Informatics and Technology Innovation, Office of Operations, Office of the Commissioner, U.S. Food and Drug Administration, Silver Spring, MD, USA*
[3]*Division of Epidemiology, U.S. Food and Drug Administration, Silver Spring, MD, USA*

ABSTRACT

Purpose Signal detection methods have been used extensively in postmarket surveillance to identify elevated risks of adverse events associated with medical products (drugs, vaccines, and devices). However, current popular disproportionality methods ignore useful information such as trends when the data are aggregated over time for signal detection.

Methods In this paper, we applied change point analysis (CPA) to trend analysis of medical products in a spontaneous adverse event reporting system. CPA was used to detect the time point at which statistical properties of a sequence of observations change over time. Two CPA approaches, change in mean and change in variance, were demonstrated by an example using neurostimulator adverse event dataset.

Results Two significant change points associated with upward trends were detected in June 2008 ($n = 20$, $p < 0.001$) and May 2011 ($n = 51$, $p = 0.003$). Further investigation confirmed battery issues and expansion of the indication for use could be possible causes for the occurrence of these change points. Two time points showed extremely low number of loss of therapy events, two cases in October 2009 and three in November 2009, which could be the result of reporting issues such as underreporting.

Conclusion As a complimentary tool to current signal detection efforts at FDA, CPA can be used to detect changes in the association between medical products and adverse events over time. Detecting these changes could be critical for public health regulation, adverse events surveillance, product recalls, and regulators' understanding of the connection between adverse events and other events regarding regulated products. © 2015 The Authors. *Pharmacoepidemiology and Drug Safety* published by John Wiley & Sons, Ltd.

KEY WORDS—change point analysis; adverse event; signal detection; pharmacoepidemiology

Received 26 September 2014; Revised 2 March 2015; Accepted 13 March 2015

INTRODUCTION

Americans rely on the Food and Drug Administration (FDA) to keep their food and medical products safe and effective. During the approval process for a medical product, such as a vaccine, drug, or medical device, manufacturers conduct rigorous analytical studies or clinical trials, and FDA carries out a thorough premarket review to evaluate the product's efficacy and safety performance. However, in their submissions to the FDA, sponsors have only tested their products on a limited number of patients from the population in which the product will ultimately be used. Therefore, it is possible that some rare adverse events from patients in the product's intended population may not be detected before the product goes to the market. In addition, in some cases, the product may change over time as newer iterations of the device are introduced, the product is used off label, or the indication for use is modified. Therefore, it is important to monitor medical products during the post-approval phase to detect emergent adverse events. To this end, FDA maintains several spontaneous adverse events reporting systems, such as FDA Adverse Event Reporting System (FAERS) for drug and biological products and Manufacturer and User Device Experience (MAUDE) for medical devices.

The spontaneous adverse event reporting systems continuously generate large volumes of data. For example, FAERS contains approximately nine million reports and currently receives approximately half a million reports per year.[1] As a result, it is not practical to manually review all reports to identify adverse event

*Correspondence to: Z. Xu, Division of Biostatistics, US FDA, 10903 New Hampshire Avenue, Silver Spring, MD, USA. E-mail: zhiheng.xu@fda.hhs.gov

†Prior postings and presentations: FDA internal presentations in 2014, Joint Statistical Meeting presentation in 2012.

safety "signals"—real changes and differences in underlying event rates. Neither is it realistic to detect subtle signals through manual review. Furthermore, there is additional difficulty in identifying signals when the total number of patients using a certain product is unknown, making the estimation of adverse event rate impossible. Thus, various data mining techniques have been implemented at FDA to detect possible safety signals, which could reveal the association between approved medical products and adverse events caused by these products. The most commonly used data mining techniques at FDA include proportional reporting ratio (PRR) and Multi-item Gamma Poisson Shrinker (MGPS). For a specific adverse event *j* and drug/device *i*, PRR is defined as the ratio of conditional probability of adverse event *j* given drug/device *i* and conditional probability of adverse event *j* given all other drugs/devices except drug/device *i*.[2] The purpose of PRR is to demonstrate the extent to which a specific adverse event is associated with that drug/device as compared with other drugs/devices. DuMouchel[3] developed the MGPS method by assuming that counts of reports containing drug *i* and adverse event *j* follow a Poisson distribution with unknown parameter λ_{ij}. A mixture of two Gamma distributions is used as the prior distribution for λ_{ij}. Five parameters are estimated from the entire data matrix, and the posterior distribution of each λ_{ij} is used to create "shrinkage" estimates, the empirical Bayes geometric mean (EBGM), which is used to rank all cell counts to determine which cells have unusually large observed counts compared with the expected counts. The lower and upper limits of a 90% confidence interval of the EBGM are denoted as EB05 and EB95, respectively. In general, safety signals will be generated if $EB05 > 2$, which means the observed count for drug/devices *i* and adverse event *j* is at least twice the expected ratio relative to all other drugs/devices and events in the database. A signal can be further refined and investigated to see if the EB05 is larger than the expected ratio of specific drugs/devices or a similar class of drugs/devices in the database. Various commercially available software programs can generate PRR and/or EBGM scores (e.g., Empirica Signal™, PVAnalyser™, SAS™, and MASE™).

Most disproportionality methods such as PRR and MGPS look at the aggregated data over time. Such methods are designed to detect a proportional increase in events for a particular drug or device as compared with a comparator set of drugs or devices. However, useful information across the timeline is lost when data are aggregated over time. Specifically, identification of trends or changes over time for a particular product may be difficult to detect.[4,5] Also, it is not always clear what constitutes the appropriate set of comparator drugs or devices. If the drug or device is the only available treatment in its class, there may not be a meaningful comparator. In other cases, there may be a wide range of similar or not-so-similar products that could be chosen to be included in a comparator set.

In this paper, we applied a time-series method, change point analysis (CPA), to trend analysis using data from the spontaneous adverse event reporting system. CPA is a powerful statistical method in determining whether a change has taken place in time series or sequences. It has been demonstrated to be an effective tool in detecting changes in different application areas such as economics, medicine, agriculture, and machine intelligence.[6–8] Recently, CPA has been introduced to public health surveillance. Kass-Hout *et al.*[9,10] applied CPA to the active syndromic surveillance data to detect changes in the incidence of emergency department visits due to daily influenza like illness during the H1N1 pandemic. To the best of our knowledge, CPA has not been used in signal detection efforts at FDA. In this study, we intended to explore CPA method as a complimentary tool in detecting safety signals from MAUDE by evaluating benefits of using CPA in detecting adverse event change points in postmarket safety surveillance. Two different CPA approaches—change in mean and change in variance—were used to investigate trends of adverse events related to a neurostimulator.

METHOD

Data source

Using the FDA MAUDE database, we retrieved adverse events from one specific neurostimulator and aggregated monthly counts for adverse events related to loss of therapy. Loss of therapy could include several types of events, including battery problems, infection, overstimulation, and so on. The data spanned from 2000 to 2012. Figure 1 illustrates monthly counts of loss of therapy during the study period.

CPA methods

The outcome measure was the monthly count of adverse event reports classified as loss of therapy. The detection of a single change point can be posed as a hypothesis test. The null hypothesis, H_0, corresponds to no change point, and the alternative hypothesis, H_a, corresponds to a single change point. The current CPA research focused on developing robust algorithms to detect multiple change points on the mean of a sequence of observation data, including binary segmentation,[11] segment neighborhoods,[12,13] and the

DOI: 10.1002/pds

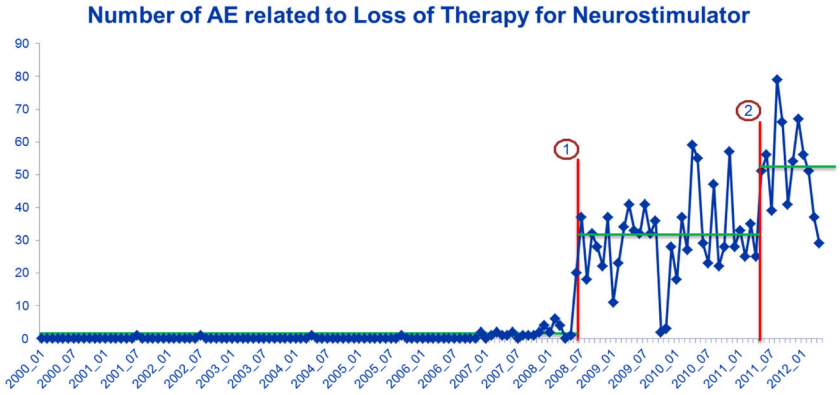

Figure 1. The time series of number of loss of therapy for neurostimulator and their detected change points. AE, adverse event

Pruned Exact Linear Time.[14] Likelihood ratio and cumulative sum (CUSUM) are two widely used test statistics in detecting changes in mean.[15] In this paper, we employed Taylor's nonparametric CPA method, which uses iterative application of CUSUM and bootstrapping methods to detect changes in time-series data.[16] This approach is based on the mean-shift model and assumes that residuals are independent and identically distributed with a mean of zero. For time-series data Y_i with $i=1, \ldots, N$, the mean-shift model is written as

$$Y_i = \mu + \varepsilon_i$$

where μ is the sample average $\mu = \sum_i Y_i/N$ and ε_i is the residual term $\varepsilon_i = Y_i - \mu$ for the ith observation. To carry out the nonparametric CPA method, we defined the CUSUMs of residuals as S_i for $i=1, \ldots, N$, where the first set $S_0 = 0$ and the remaining sets were calculated as $S_i = S_{i-1} + \varepsilon_i$ for $i=1, \ldots, N$. Note that by construction, because we were subtracting the overall mean, $S_N = 0$ as well. If there was no change point, the time series is stationary, and thus a permuted sample of the residuals can be used to construct an instance of the CUSUMs under the null. Repeated permutation samples can be used to provide a null distribution for a test statistic constructed from the CUSUMs.

A potential change point in an interval was identified at location m by searching for the maximum absolute CUSUM of residuals, where $|S_m| = \max_{i=0,\ldots,N} |S_i|$. As a statistic, we used the maximum absolute CUSUM difference within a given interval $S_{\text{diff}} = S_{\max} - S_{\min}$, where $S_{\max} = \max_i S_i$ and $S_{\min} = \min_i S_i$. On the other hand, when the CUSUM of residuals were plotted, a sudden change in direction of the CUSUM indicated a sudden shift or change in the average, and the place where sudden change occurred was defined as change point. The distribution of 1000 S_{diff} was used to determine the p-value for the change point as the percentage of S_{diff} values, which were greater than S^0_{diff} from original time-series data.[16]

In addition to the changes in mean approach, we used another CPA method for detecting changes in variability. This method was motivated by potential non-stationarity of variability of the data where variance was smaller in some sections but larger in other sections.[15] Similar to the changes in mean CUSUM approach, change in variance can be implemented using the sum square error (SSE) approach, which was meaningful in operation, simple in calculation, and useful for testing significance. Let $SSE(m)$ be defined as

$$SSE(m) = SSE_1 + SSE_2 = \sum_{i=1}^{m} \left(X_i - \overline{X}_1\right)^2 + \sum_{i=m+1}^{n} \left(X_i - \overline{X}_2\right)^2$$

where $\overline{X}_1 = \dfrac{\sum_{i=1}^{m} X_i}{m}, \quad \overline{X}_2 = \dfrac{\sum_{i=m+1}^{n} X_i}{n-m}$

From the analysis of variance, it was known that the sum of the squared distances of points on a line from their mean can be partitioned, when the points were classified into two groups, 1 to m and $m+1$ to N, into two within-group sums of squares SSE_1 and SSE_2 and a between-groups sum of squares SSE.[11] The change point was defined at the value of m that minimizes $SSE(m)$, the sum of the two within-group sums of squares. This can be thought of as a modified application of the k-means clustering algorithm, which was used to partition n observations into k clusters with each observation belonging to the cluster with the nearest mean. In this application, the clusters were restricted to retain the time-series nature of the data.

The open-source software R has a package called *changepoint*, which provides a choice of different CPA algorithms in detecting changes in mean and variances. We used *cpt.var* function in the R package *changepoint* to implement the change in variance approach. For the change in mean CUSUM CPA method, we have developed publically available codes in R, SAS, and STATA format. Those codes can be downloaded from our open-access collaboration website for CPA at https://sites.google.com/site/changepointanalysis.

RESULTS

The CUSUM plot for the sample neurostimulator data is shown in Figure 2. Figure 2(a) shows the CUSUM direction for entire time-series data, while Figure 2(b) and (c) displays the CUSUM trend prior and post the first detected change point on June 2008.

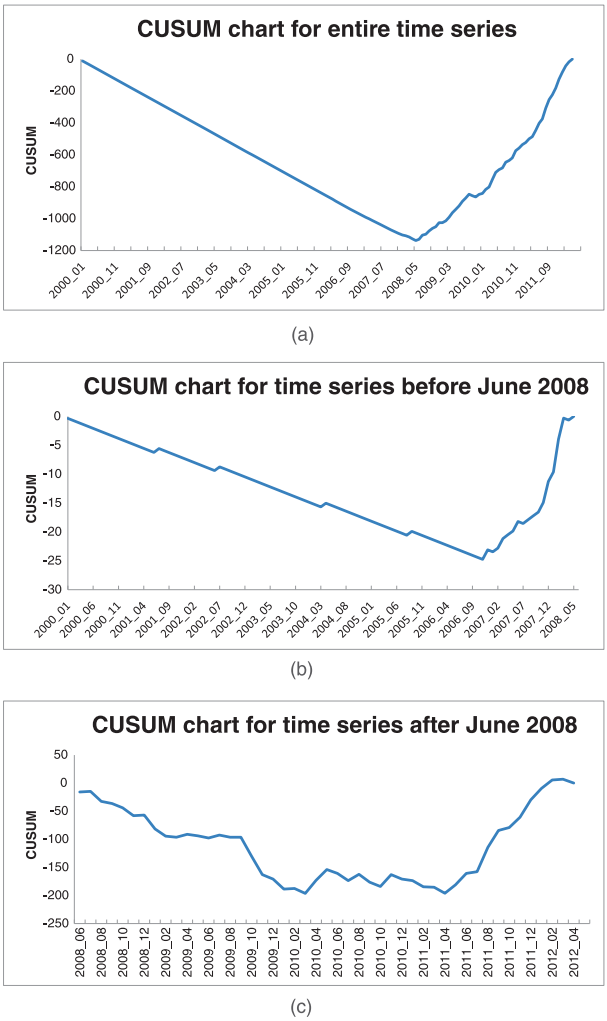

Figure 2. (a) Cumulative sum (CUSUM) plot for the sample neurostimulator data; (b) CUSUM plot for the sample neurostimulator data prior to first change point at June 2008; and (c) CUSUM plot for the sample neurostimulator data after first change point at June 2008

Change points detected using CUSUM for the sample neurostimulator data along with their significance levels are listed in Table 1. A p-value of 0.05 was used as the cutoff to screen significant change points. The first candidate change point occurred in June 2008, with a permutation test p-value of <0.001.

Figure 1 shows how the data series were split based on the change points. The June 2008 change point is displayed as symbol "1" in Figure 1. Then the data were split into two segments in June 2008, and CPA was implemented independently on each of the resulting segments. The second significant change point occurred in May 2011, which is displayed as symbol "2" in Figure 1. This segment was further split in May 2011 change point, and CPA was applied, but no further significant change point was detected. For the left segment (prior to June 2008), CPA picked up three additional significant change points in 2006 and 2007.

Because the variability of the data showed two different patterns before and after June 2008, we applied a change in variance CPA approach to detect any significant change points caused by the variability of the data (Table 2). The change in variance CPA method detected two change points, June 2008, which was also detected by change in mean method, and April 2011, which is 1 month before the change point detected by change in mean method.

DISCUSSION

Change point analysis is a very useful tool for detecting significant changes in the means or variances of a sequence of observed data. For medical products,

Table 1. Change points using cumulative sum method based on changes in mean

Detection order	Change point	Count in change point month	p-value
1	June 2008	20	<0.001
2	May 2011	51	0.003
3	December 2006	2	<0.001
4	November 2007	2	0.005
5	February 2007	1	<0.001

Table 2. Change points using change point analysis method based on changes in variance

Change point	Count in change point month
June 2008	20
April 2011	25

DOI: 10.1002/pds

the detected change points may provide valuable information for postmarket surveillance. In the neurostimulator example we used in this paper, two significant change points were detected—June 2008 and May 2011. A report based on retrospective analysis shows battery, and other device failure issues also occurred during June 2008. In May 2011, an extended indication for use for this device was approved. As a larger population used this device, more adverse events were reported. It is less intuitive as to observe significant change points before 2008 because the number of reports was relatively small and, for most months, there was no report. One of the limitations of the CUSUM approach is that it is based on the identification of change in mean, and thus, its performance could be affected by the actual numerical value of data. The counts of two losses of therapy could be significant change points if there are no events nearby (i.e., $n=2$ for December 2006 and November 2007). However, although statistically significant, a count of two may not be sufficient in providing meaningful clinical implication to call for further investigation. Considering there are millions of adverse event reports in FAERS and MAUDE, it would be ideal to reduce the number of change points, which require further action.

We also observed several low numbers of reports from October 2009 to January 2010, for example, two for October 2009 and three for November 2009. Because they existed in the elevated reporting period, the unusual pattern may be due to reporting issues such as underreporting.

Underreporting of events is a significant problem in the spontaneous adverse event reporting system. Reporting of adverse events and medication errors by healthcare professionals and consumers is voluntary in the USA. As a result, a significant underreporting of adverse events occurs.[17] FDA receives some adverse event, and medication error reports directly from healthcare professionals (e.g., physicians, pharmacists, and nurses) and consumers (e.g., patients, family members, and lawyers). Healthcare professionals and consumers may also report adverse events and/or medication errors to product manufacturers. If a manufacturer receives an adverse event report, it is required to send the report to FDA as specified by regulations. Reports received directly by FDA and reports from manufacturers are entered into FAERS or MAUDE. Underreporting may vary according to the type of product, the seriousness of an event, the population using the product, the product's time on the market, and other factors. It has been estimated that 94% of adverse drug reactions go undetected by spontaneous reporting systems.[17]

In addition to underreporting, overreporting of events could also be problematic. A relative increase of reporting for a particular event or syndrome of events may be stimulated by publicity or litigation.[18] However, such inflated reporting could lead to a biased estimate of safety signals. Factors that can cause bias include newly published safety alerts and recalls, new regulations, reporting incentives, or deterrents. Furthermore, when the report is from patients who are using multiple medical products, it is hard to correctly identify which medical products caused the adverse event.

Both change in mean and change in variance CPA approaches showed similar change points in the neurostimulator adverse event dataset. April 2011 can be considered the last time point before counts of loss of therapy went up in May 2011. The CUSUM method treated the beginning of new segment as the change point, while the change in variance CPA approach used the last time point before the change occurred. We would argue that both change points (April and May 2011) represent the same change that occurred in the data. One difference between the two approaches is that change in mean detected multiple significant change points in 2006 and 2007 while change in variance did not. Compared with change in mean, change in variance is less sensitive to the actual numerical value of the data and could be more robust to outliers.

The approaches presented in this paper provide complimentary analyses to findings from disproportionality-based methods. Instead of aggregating data to detect differences in relative reporting rates between products, CPA methods focus on detecting changes within a product over time. Both types of analysis can provide signals that would prompt further investigation of potential problems. CPA should be used as the starting point, not the end point to investigate these changes. In some situations, the actual events may start even before the detected change points. In order to determine the underlying cause of those changes, multiple data sources may be used for the investigation. As the volume of time-series data increases, there is a growing need to maintain situational awareness and be able to efficiently and accurately estimate the location of multiple change points. As a complimentary tool to current signal detection efforts at FDA, CPA can be applied to detect changes in the association between medical products (drugs, vaccines, and devices) and adverse events over time. Detecting these changes could be critical for public health regulation, adverse events surveillance, product recalls, and regulators' understanding of the connection between adverse events and other events regarding regulated products.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

KEY POINTS

- Signal detection using disproportionality methods may ignore useful information when the data are aggregated over time for signal detection.
- Change point analysis (CPA), a time-series analysis tool, allows the estimation of the point at which statistical properties of a sequence of observations change.
- CPA can be used to detect changes in mean or in variance.
- CPA can be applied to detect changes in the association between medical products (drugs, vaccines, and devices) and adverse events over time.
- CPA can be a complimentary tool to current signal detection efforts at FDA.

ETHICS STATEMENT

This research analyzes the publicly available MAUDE data and does not require ethical approval.

ACKNOWLEDGEMENTS

We are grateful for Dr Lorna Ye who helped with the proofreading and language editing.

DISCLAIMER

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the FDA.

REFERENCES

1. Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. Performance of pharmacovigilance signal-detection algorithms for the FDA Adverse Event Reporting System. *Clin Pharmacol Ther* 2013; **93**(6): 539–546. doi:10.1038/clpt.2013.24.
2. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf* 2001; **10**(6): 483–6.
3. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *The Am Stat* 1999; **53**(3): 177–190.
4. Lao CS, Kessler LG, Gross TP. Proposed statistical methods for signal detection of adverse medical device events. *Drug Inf J* 1998; **32**: 183–191.
5. Lao CS. Statistical issues involved in medical device postmarketing surveillance. *Drug Inf J* 2000; **34**: 483–493.
6. Finney DJ. Field sampling for the estimation of wireworm population. *Biometrics* 1946; **2**(1): 1–7.
7. Hensen B. The new econometrics of structural change: dating changes in U.S. labor productivity. *J Econ Perspect* 2001; **15**: 117–28.
8. Erdman C, Emerson J. A fast Bayesian change point analysis for the segmentation of microarray data. *Bioinformatics* 2008; **24**(9): 2143–48.
9. Kass-Hout TA, Xu Z, McMurray P, *et al.*. Application of change point analysis to daily influenza-like illness (ILI) emergency department visits. *J Am Med Inform Assoc* 2012. doi:10.1136/amiajnl-2011-000793.
10. Kass-Hout TA, Xu Z. Change point analysis. Retrieved from https://sites.google.com/site/changepointanalysis on Feb. 28, 2014.
11. Edwards AWF, Cavalli-Sforza LL. A method for cluster analysis. *Biometrics* 1965; **21**: 362–375.
12. Auger IE, Lawrence CE. Algorithms for the optimal identification of segment neighborhoods. *Bull Math Biol* 1989; **51**(1): 39–54.
13. Bai J, Perron P. Estimating and testing linear models with multiple structural changes. *Econometrica* 1998; **66**(1): 47–78.
14. Killick R, Fearnhead P, Eckley IA. Optimal detection of changepoints with a linear computational cost. *J Am Stat Assoc* 2012; **107**(500): 1590–1598.
15. Killick R, Eckley I. Changepoint: an R package for changepoint analysis. Retrieved from http://www.lancs.ac.uk/~killick/Pub/KillickEckley2011.pdf on Feb. 28, 2014.
16. Taylor W. Change-point analysis: a powerful new tool for detecting changes. Available from: http://www.variation.com/anonftp/pub/changepoint.pdf on Feb. 28, 2014.
17. Hazell L, Shakir SA. Under-reporting of adverse drug reactions: a systematic review. *Drug Saf* 2006; **29**(5): 385–96.
18. McAdams M, Staffa J, Dal pan G. Estimating the extent of reporting to FDA: a case study of statin-associated rhabdomyolysis. *Pharmacoepidemiol Drug Saf* 2008; **17**(3): 229–39.