

Occupational Safety Surveillance Using a Statistical Monitoring Approach

Anna Kristine Schuh

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Industrial and Systems Engineering

Jaime A. Camelio, Chair
Brian M. Kleiner
Eileen M. Van Aken
William H. Woodall

April 1, 2013
Blacksburg, VA

Keywords: Occupational Safety, Accident Surveillance, Accident Frequency, Accident Severity,
Statistical Monitoring, Control Charts, Cumulative Sum (CUSUM), Poisson CUSUM,
Exponential CUSUM, Aggregated Data

© 2013, Anna Kristine Schuh

Occupational Safety Surveillance Using a Statistical Monitoring Approach

Anna Kristine Schuh

ABSTRACT

When unsafe conditions arise in a workplace, they may result in employee accidents and fatalities. However, if these problems are detected early, new hazard controls and safety initiatives can be introduced in order to actively reduce or prevent the occurrence of these events. Unfortunately, many safety systems currently monitor and report data that has been aggregated over long time periods, making it difficult to realize and respond to pattern shifts in a timely manner.

When monitoring a process over time, a commonly used tool is statistical process control charting. Traditionally used in manufacturing, control charts indicate a deviation from historically “normal” or “in-control” behavior and have become increasingly common in healthcare and public health monitoring. This dissertation studies the use of control charts to monitor the frequency of occupational safety incidents, with the overarching goal of investigating the effects of data aggregation on the detection performance of these charts.

Specifically, this dissertation 1) qualitatively establishes the need for more frequent monitoring of safety incidents; 2) investigates the comparative performance of control charts with aggregated and non-aggregated data for the detection of increased accident frequency, using a case study with data from an industrial partner; 3) more generally compares the performance of these charts for a Poisson process with a range of simulated process shifts; and 4) discusses the potential future challenges of including accident severity in quantitative safety monitoring systems. The comprehensive results indicate that lower degrees of data aggregation are preferred, and suggestions for better data collection and employee communication practices are offered to aid the transition for companies.

DEDICATION

To Mom and Dad,

this is the product of a lifetime of encouragement.

ACKNOWLEDGEMENTS

First, I will thank Him.

But then I must thank you.

If you are Dr. Jaime Camelio:

...thank you for challenging me, believing in me, and celebrating every success.

If you are Dr. Bill Woodall:

...thank you for your meticulous commitment to me, our work together, and this research topic.

If you are Dr. Brian Kleiner:

...thank you for sharing both your invaluable expertise and your sense of humor.

If you are Dr. Eileen Van Aken:

...thank you for your enthusiastic support and feedback throughout this work.

If you are a past or present member of the IMAS/CIBM lab:
(especially, but not limited to, Jeremy Rickli, Lee Wells, Dr. Fadel Megahed, Gregory Purdy,
and Cory Niziolek):

...thank you for constantly teaching me new things about research, engineering, and life.

If you are Michael Bylund:

...thank you for your dedication to our industry project.

If you are a past or present fellow ISE graduate student:
(especially, but not limited to, Dr. Adria Markowski, Nevin Mutlu, Esra Ağca, Katharine Mann,
Dr. Lora Cavuoto, Stephanie Alpert, and Leily Farrokhvar)

...thank you for sharing this experience with me as both friends and colleagues.

If you are my friends in Blacksburg:

...thank you for all the shopping dates, happy hours, football games, parties, and laughs.

If you are my friends and family in Minnesota:

...thank you for loving me.

You all mean more to me than you know.

CONTENTS

1. Introduction	
1.1 Motivation.....	1
1.2 Research Objectives.....	2
1.3 Dissertation Preview.....	2
2. Traditional Performance Measurement in Construction Safety: A Case for Active Safety Monitoring	
2.1 Abstract.....	4
2.2 Introduction.....	4
2.3 Methodology.....	6
2.4 Results and Discussion.....	14
2.5 Conclusions and Future Work.....	17
2.6 References.....	18
3. Control Charts for Accident Frequency: A Motivation for Real-Time Occupational Safety Monitoring	
3.1 Abstract.....	21
3.2 Introduction.....	21
3.3 Cumulative Sum Control Charts for Monitoring Incident Data.....	23
3.4 Application to Industrial Accident Data.....	27
3.5 Conclusions.....	34
3.6 Current Limitations and Opportunities for Future Work.....	35
3.7 References.....	37
4. The Effect of Aggregating Data When Monitoring a Poisson Process	
4.1 Abstract.....	41
4.2 Introduction.....	41
4.3 Poisson and Exponential CUSUM Charts	44
4.4 Relative Performance of Poisson and Exponential CUSUM Charts.....	46
4.5 Conclusions.....	60
4.6 References.....	61
5. Including Accident Severity in Statistical Monitoring Systems	
5.1 Abstract.....	64
5.2 Introduction.....	64
5.3 Statistical Monitoring Background.....	65
5.4 Severity Metrics.....	67
5.5 Safety Management System Methods for Accident Severity.....	71
5.6 Conclusions and Future Work.....	75
5.7 References.....	76

6. Conclusions and Future Work	
6.1 Contributions.....	80
6.2 Data Collection and Analysis.....	82
6.3 Control Charting Considerations.....	83
6.4 Value Communication and Usability.....	85
References Not Included in Chapters 2-5.....	87
Appendix A: Institute of Industrial Engineers Copyright Permission.....	89

LIST OF FIGURES

2.1	Industrial hygiene Hierarchy of Hazard Controls (adapted from CDC, 2012).....	6
2.2	The Safety Decision Hierarchy (adapted from Manuele, 2005).....	7
2.3	Proposed cycle for implementation of the continuous improvement methodology.....	14
2.4	Word cloud data visualization of incident types prevalent in industrial data.....	15
2.5	Historical Safety Incident Indicator values from Firm X with current performance measurements (a) and expected future Safety Incident Indicator values using the proposed monitoring process (b).....	16
3.1	Year 2 Poisson CUSUM control charts for aggregated accident counts, a) quarterly; b) monthly; c) semi-monthly; d) weekly; e) every 3 days; f) daily.....	29
3.2	Year 2 Time-between-events CUSUM control chart.....	30
3.3	Simulated future performance of a) the daily Poisson CUSUM chart and b) the time-between-events CUSUM chart with a thirty day period for control implementation.....	32
3.4	Simulated future performance of a) the daily Poisson CUSUM chart and b) the TBE CUSUM chart with a 10-day period for control.....	33
3.5	Simulated future performance of a) the daily Poisson CUSUM chart and b) the TBE CUSUM chart with adjusted starting and re-starting values.....	34
4.1	Zero-state (a) and steady-state (b) scenarios for monitoring with aggregated Poisson counts.....	50
4.2	Zero-state (a) and steady-state (b) scenarios monitoring with exponential interarrival times.....	51
4.3	Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency, based on data from Table 4.4.....	55
4.4	Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency, based on decreases to $\lambda=0.2-0.85$ from Table 4.6.....	59
5.1	Exponential CUSUM control chart for monitoring accident frequency.....	66
5.2	Risk Assessment Matrix.....	67
5.3	Total lost work days resulting from a safety incident, count updated monthly.....	68

5.4 Event management points of interest with a control charting system..... 75

LIST OF TABLES

2.1	Mapping of current performance measurements to perceived levels of achievement on the Hierarchy of Hazard Controls.....	10
2.2	Mapping of the proposed active monitoring methodology to perceived levels of achievement on the Hierarchy of Hazard Controls.....	16
4.1	Gan (1994) zero-state ATS for optimal (in detecting shift to $\lambda = 0.5$) Poisson CUSUM and exponential CUSUM control charts.....	46
4.2	Steady-state average time to signal for optimal (in detecting a sustained shift to $\lambda = 0.5$) Poisson and exponential CUSUM control charts with respect to an increase in the rate of occurrences of events.....	47
4.3	Steady-state performance for optimal (in detecting shift to $\lambda = 0.5$) Poisson and exponential CUSUM control charts with respect to an increase in the rate of occurrences of events.....	53
4.4	Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency.....	54
4.5	Steady-state performance for optimal (in detecting shift to $\lambda = 0.75$) Poisson and exponential CUSUM control charts with respect to an increase in the rate of occurrences of events.....	57
4.6	Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency.....	58
4.7	Steady-state performance for Poisson CUSUM charts (POIS30 from Table 4.4, and POIS365 from Table 4.6) and Shewhart control charts, with respect to an increase in the rate of occurrences of events.....	60
5.1	Safety Management System approaches applied to control charting of severe accidents.....	74

1. Introduction

1.1 Motivation

The United Nation's International Labour Organization (ILO) estimates that there are annually 337 million occupational accidents and 160 million occupational diseases, resulting in a total of 2.3 million worker deaths per year (Niu, 2010). In the U.S., approximately 4,600 fatal work injuries occurred in 2011; 3.5 per 100,000 employees were killed on the job, and the average non-fatal injury resulted in 8 days away from work (Occupational Safety and Health Administration, OSHA, 2013; Bureau of Labor Statistics, BLS, 2013). Comparing the number of worker fatalities in the U.S. to that in other industrialized countries for the construction industry in particular suggests poor performance globally. In 2005, the U.S. reported 11.1 deaths from injuries per 100,000 construction workers (CPWR, 2008). Compared to much lower rates in other countries the same year (Finland, 9.8; Norway, 7.0; Germany, 6.0; Australia, 5.9; Canada, 5.0; Switzerland, 4.8; and Sweden 4.4), it is clear why the National Institute of Occupational Safety and Health (NIOSH), OSHA and many individual firms are concerned about improving safety performance in the U.S.

In order to reduce the rates of occupational fatalities and nonfatal injuries, it is important to measure, monitor, and report safety performance over time, and act accordingly. However, current safety performance measurement systems in the U.S. rely predominantly on collecting aggregated data about past accidents and attempting to respond to reports based on this aggregated data. These indicators are known as reactive or "lagging" indicators, since current and future responses are based on reactions to past experiences. In some cases, such as reports released by BLS based on data collected by the Occupational Safety and Health Administration (OSHA), the reports are only released on an annual basis and data from a given year is not reported until September of the following year (BLS, 2013). This creates at least a nine-month delay in the reporting of data needed by industries to make quantitative improvement decisions. In this way, annual reporting makes it nearly impossible to use this data for the purpose of improving behaviors in occupational safety (Mohamed, 2003). NIOSH recognizes the need for better safety surveillance and has made it a priority in their strategic goals (2011).

1.2 Research Objectives

This dissertation will posit that better safety will result from the transition from aggregated to real-time accident reporting, which will in turn aid transitions from lagging to leading indicators and from reactive to proactive hazard controls. Focusing on the execution of this first transition from aggregated to real-time monitoring, the objectives of this work are to:

- Demonstrate the value of real-time accident monitoring through a continuous improvement methodology using real historical accident data;
- Emphasize this value quantitatively through statistical process control charting methods using real historical accident data;
- Investigate the effects of data aggregation when monitoring a Poisson process, which can be assumed to describe accident frequency; and
- Evaluate the opportunities and limitations of expanding these monitoring methods to include accident severity.

1.3 Dissertation Preview

This dissertation is comprised of six chapters. The current introductory chapter has motivated the four manuscripts, which follow in Chapters 2-5. Chapter 6 offers a comprehensive discussion with conclusions and suggestions for future work.

Chapter 2 of this dissertation suggests a new methodology for identifying the accident types in most need of new hazard controls. This new methodology incorporates elements of the industrial hygiene Hierarchy of Controls, the Plan-Do-Check-Act process improvement cycle, and the automobile industry's "2mm Program" quality improvement initiative. The proposed methodology is used to identify and monitor safety data classes which highlight the most common accident types. The process is validated with real data from a manufacturing partner. The work in this chapter has been submitted for publication and is a collaborative work with Dr. Jaime Camelio and Dr. Brian Kleiner.

However, it may be desirable to monitor accident frequency in a more quantitative manner. A common method for process monitoring over time is the use of statistical process control charts. These charts have been traditionally used in manufacturing settings to monitor the occurrence of

nonconforming parts, and have more recently been used in public health applications. Chapter 3 provides an initial investigation into the use of control charts to monitor accident frequency. The comparative performance of Poisson CUSUM charts for aggregated data and exponential CUSUM charts for non-aggregated data is presented. In order to demonstrate the applicability of the proposed methods, an example using real data from an industrial partner is used for this comparison. The manuscript given in this chapter has been accepted for publication in the *International Journal of Injury Control and Safety Promotion*, and is a joint effort with Dr. Jaime Camelio and Dr. William Woodall.

Chapter 4 more generally investigates the effect of data aggregation for a Poisson process over time. Simulated data is used in order to consider these effects for a wide range of process shifts. The paper presented in this chapter has been accepted for publication in the *Journal of Quality Technology* with co-authors Dr. William Woodall and Dr. Jaime Camelio.

Finally, when considering safety events it may be desirable to monitor incidents according to their severity and not only their frequency, to better prioritize the implementation of new safety practices. Chapter 5 presents an initial discussion about using control charts to quantitatively monitor accident severity. Current severity metrics are reviewed and reactive, proactive, and predictive safety management techniques are discussed. This paper, written with Dr. Jaime Camelio, will appear in the *2013 Proceedings of the Industrial and Systems Engineering Research Conference*.

Chapter 6 provides a comprehensive discussion and conclusion of overall results. Suggestions for future work focus on better safety data collection, extensions for similar statistical monitoring applications, and the collaboration efforts required to connect the fields of occupational safety and statistical monitoring.

2. Traditional Performance Measurement in Construction Safety: A Case for Active Safety Monitoring

2.1 Abstract

This paper provides a review of the main limitations in the performance measurements currently used to assess occupational safety in the U.S. construction industry. Safety has traditionally been a difficult performance dimension to measure quantitatively. Predictive approaches are virtually non-existent, and current performance measurement systems do not provide enough information to quickly implement effective hazard control. Elements from the Hierarchy of Hazard Controls, the Plan-Do-Check-Act cycle, and the automobile industry's quality improvement *2mm* Program are combined into a prospective active monitoring methodology. The new methodology focuses on quickly reducing the highest priority incidents, including near misses. The expected performance of the new system is validated using data provided by an industrial partner. Results are discussed in terms of the expected contribution of the new system to the implementation of elimination and substitution hazard controls.

2.2 Introduction

The U.S. construction industry continues to yield high rates of fatal and nonfatal occupational injuries. In 2005, the United States had a rate of 11.1 deaths from injuries per 100,000 construction workers (CPWR, 2008). When compared to much lower rates in other countries the same year (Finland, 9.8; Norway, 7.0; Germany, 6.0; Australia, 5.9; Canada, 5.0; Switzerland, 4.8; and Sweden 4.4), it is clear why NIOSH, OSHA and many construction firms are concerned.

As learned in other industries, it is important to measure, monitor, and report safety performance over time, and act accordingly in order to reduce the rates of occupational fatalities and nonfatal injuries. However, current safety performance measurement systems in the U.S. construction industry rely predominantly on collecting aggregated data about past accidents and attempting to respond to reports based on this aggregated data. These indicators are known as reactive or "lagging" indicators, since current and future responses are based on reactions to past incidents. In some cases, such as reports released by the Bureau of Labor Statistics (BLS) based on data collected by the Occupational Safety and Health Administration (OSHA), the

reports are only released on an annual basis and data from a given year is not reported by BLS until September of the following year (BLS, 2013). Annual reporting makes it nearly impossible to use this data for the purpose of improving behaviors in construction environments (Mohamed, 2003). Hinze et al. (2013) provided an extensive review of lagging and leading indicators.

In this paper, the shortcomings of current performance surveillance methods are illustrated through a mapping to the Hierarchy of Hazard Controls. The hierarchy is a standard process promoted by researchers and federal regulatory bodies and used by safety managers to eliminate or mitigate hazards. The Hierarchy of Controls prioritizes methods of reducing hazards, ranging from most effective (total elimination of risk) to least effective (wearing of personal protective equipment). However, actual implementation of the most effective levels of the hierarchy requires proactivity and enough knowledge about incidents to make appropriate process changes (Manuele, 2005). It is posited that current performance measurements do not provide this requisite knowledge, so a new process for active safety monitoring is proposed to address these shortcomings.

Our proposed monitoring process introduces a new Safety Incident Indicator, which is monitored much more regularly than typical measurement systems, so that trends in the frequency of predominant incident types will be more evident. Furthermore, data about “near misses,” or incidents, which nearly cause a reportable injury, can be easily integrated our new monitoring process. In this way, the Safety Incident Indicator becomes a proactive, leading indicator system which can signal unsafe conditions before injury or fatality, thus preventing traumatic or large scale events.

The remainder of this paper is organized as follows: Section 2.3 further discusses the Hierarchy of Controls and presents the difficulties in achieving the effective levels with current performance measurement systems. Section 2.4 proposes the new Safety Incident Indicator process and presents an example of expected performance. Section 2.5 presents the conclusions and future work.

2.3 Methodology

2.3.1 Hierarchy of Hazard Controls

The Hierarchy of Hazard Controls provides a prioritization process for hazard elimination. Fig. 2.1 is based on the version of the Hierarchy provided by the Centers for Disease Control and Prevention (CDC, 2012). Many similar versions of this hierarchy have been used in practice (Barnett and Brickman, 1985) with levels typically ranging from 4-5, Administrative Controls and Personal Protective Equipment. The U.S. Department of Defense uses a comparable “system safety design order of precedence” (US Dept. of Defense, 2000). Some representations use a triangle or an upside-down triangle to depict the decreasing level of effectiveness.

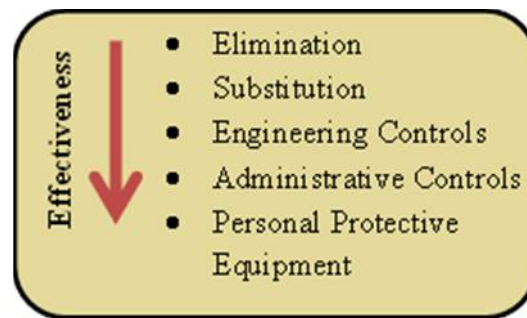


Figure 2.1: Industrial hygiene Hierarchy of Hazard Controls (adapted from CDC, 2012)

The elimination and substitution of present hazards are the most effective controls, and should be considered first; engineering controls place a barrier between the worker and the hazard that remains independent of worker interactions; administrative controls and personal protective equipment are the least effective controls, as they rely on worker behavior and adherence to policies. The most effective controls are usually system design-related and the most difficult and costly to implement in the short term (CDC, 2012), but in the long term they are much more effective (Manuele, 2005).

Manuele (2005) suggests a Safety Decision Hierarchy in which the Hierarchy of Controls is incorporated into a problem-solving process. To implement this process, one would identify hazards and assess risks, consider the controls in the hierarchy in order of decreasing effectiveness, decide on a control and take action, and finally measure the effectiveness of the control and re-evaluate as needed. This process is shown in Fig. 2.2.

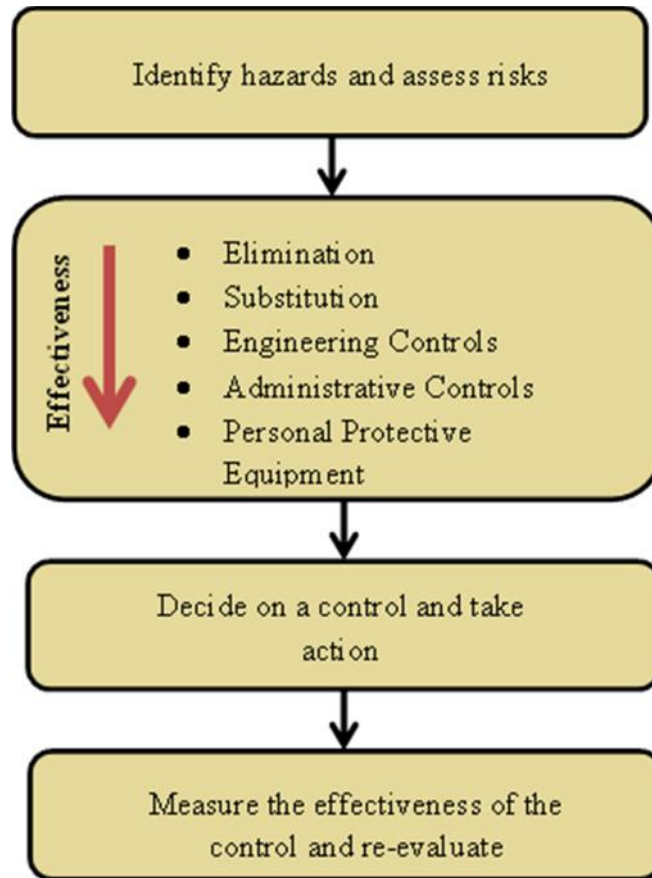


Figure 2.2.: The Safety Decision Hierarchy (adapted from Manuele, 2005)

However, current performance measurement systems in construction safety do not provide enough information about individual incidents to successfully identify hazards and assess risks, since they mostly rely on past aggregated data. Therefore, the most effective controls listed in the hierarchy will be even more difficult to implement, since hazards (and therefore precisely what to eliminate or substitute) may not be easily identified.

Next, the current commonly used safety performance measurement systems are discussed with an emphasis on how they can be used and evaluated based on the Hierarchy of Hazard Controls.

2.3.2 Current Performance Measurement Systems

Currently, construction companies which employ more than eleven workers are required by law to report each “incident” to OSHA through documentation using a standardized OSHA-300 reporting form (OSHA, 2013). OSHA collects this injury and fatality data and BLS prepares an

annual report based on these compilations. These annual reports are publically accessible on the BLS website and include tables that display the total number of workers reportedly affected by fatal and non-fatal injuries, organized by industry and type of injury (BLS, 2013). Individual OSHA-300 reports require basic information about the incident: employee name, job title, incident date, location, short incident description, incident classification (death, days away from work, restricted days, other), number of days away or in restriction, and the type of injury or illness.

Since the OSHA-300 log requires brief descriptions of incidents, and some of these brief descriptions are available to the public on a weekly basis for incidents involving fatalities (OSHA, 2013), it is possible that enough information could be present to implement effective engineering controls. However, much text and data mining would need to take place before specific hazards could be targeted for substitution or elimination.

The commonly used performance measurements in construction safety rely heavily on the numbers of nonfatal and fatal injuries recorded on the OSHA-300 log. The Accident, Incident, and Experience Modification Rates are discussed in the following sections. Some proactive, typically larger, safety-minded contractors do collect near miss and leading indicators, but these are usually time intensive and are often tied to punitive behavior-based systems.

2.3.2.1 Accident and Incident Rates

Using OSHA-300 logs, the accident rate is computed based on the total number of accidents in a year experienced by an individual company; the respective incident rate is found using more in-depth recorded data, including lost time, the number of fatalities, and the number of non-fatal illnesses and injuries in a year (Ng et al., 2005). The OSHA Recordable Incidence Rate has been touted as the best indicator of performance because it remains relatively constant over time and it is easily calculated (Tam and Fung, 1998). However, the use of accident and incident rates as the only performance metrics has been criticized for discouraging accurate reporting (Mohamed, 2003; Ng et al., 2005). If the only measure of safety performance is the lagging percentage of accidents, there is little incentive for construction managers to be diligent about recording minor incidents.

If a high accident or incident rate is observed, it is assumed that some degree of investigation into the incidents accounting for these high rates will take place. However, since incidents will only be investigated based on aggregated reports from at least a year prior, it is likely that only general information will be gleaned. This information may lead to new policies involving personal protective equipment and administrative controls, but not enough information about the causes of these incidents will be retained to implement any engineering controls or substitution or elimination of hazards, due to the lagging and static aspects of this measurement.

2.3.2.2 Experience Modification Rate

The Experience Modification Rate (EMR) is commonly used in the US construction industry to determine insurance premiums (Hoonakker et al., 2005). Developed by the National Council on Compensation Insurance (NCCI), this measure aims to minimize future losses for the insurance companies while providing an incentive to firms with good safety records (Everett and Johnson, 1995). The formula provides a ratio between a company's actual number of accidents experienced and its expected number of accidents, based on accident rates during the past three years. The average ratio should be equal to one; a number higher than one is considered poor and will indicate a higher insurance premium for the respective company (Ng et al., 2005). This measure is convenient because it is directly related to operational costs and is normalized for variables across the industry (company size, major operations, etc.). While the EMR is essentially a ratio between easily measured values, the precise formula is quite complex and involves several variables (Everett and Johnson, 1995).

Unfortunately, there are discrepancies in the EMR formulae used by individual companies. The standard EMR formula is only used in 32 states. In 13 other states, the EMR is calculated by insurance firms using similar equations. In the remaining five states, the state government determines the EMR equation (Everett and Johnson, 1995). Because some companies may use the standard formula and others may use a simpler modification (Ng et al., 2005), national comparisons of 'safety' across multiple companies based on EMR may not be reliable or valid.

As with the accident and incident rates, any information about incidents causing high EMR values will be garnered based on past data and root-cause investigation will be difficult. Again,

PPE and administrative controls will be the only controls likely to be developed based on monitoring of EMR values.

2.3.2.3 How current performance measurements map to the Hierarchy of Controls

Table 2.1 reflects the controls believed to be implementable with each of the current metrics, as discussed above.

	PPE	Administrative Controls	Engineering Controls	Substitution	Elimination
OSHA and BLS	X	X	X		
Incident Rate	X	X			
EMR	X	X			

Table 2.1: Mapping of current performance measurements to perceived levels of achievement on the Hierarchy of Hazard Controls

In the next section, a proposed active safety monitoring process is introduced. This new process addresses the most effective levels of the Hierarchy of Hazard Controls by providing more information regarding the most common accident types, doing so in a more timely manner, and allowing for the inclusion of near miss data for more proactive implementation of controls.

2.3.3 Proposed performance measurement system

From Table 2.1, it is clear that a system is needed to provide more timely and detailed incident data so that the most effective hazard controls can be implemented. Many researchers have successfully conducted root-cause analyses of incidents using aggregated OSHA reports (Hinze and Russell, 1995; Hinze et al., 1998; Huang and Hinze, 2003; and Beavers et al., 2006). However, when using aggregated data, there is an inherent time lag between the occurrence of the trend and the recognition and investigation of that trend. For instance, in their 2006 publication, Beavers et al. provided valuable insight about the causes of crane-related accidents in the U.S. construction industry using OSHA data collected between 1997 and 2003 – but this unfortunately means that the causes of the 1997 accidents were not published until nine years later. The use of timely monitoring will allow both companies and national safety organizations to recognize and investigate incident trends, implement more effective hazard controls, reduce incidents, and lower costs.

This section suggests a new continuous improvement methodology to address the shortcomings in the surveillance of aggregated data.

2.3.3.1 Continuous improvement methodologies for safety

Continuous improvement programs have been widely used in corporations to systematically work toward company goals with real-time analysis of success (Stevenson, 2009). Popular corporate efficiency initiatives such as lean manufacturing, Six Sigma, and Kaizen events, all incorporate principles of continuous improvement, and are historically focused on quality and productivity. The development of a continuous improvement methodology for safety will allow systematic measurement of improvement and introduce a near real-time monitoring component.

In his 1986 work, *Out of the Crisis*, Deming introduced the Plan-Do-Check-Act cycle for continuous improvement of quality. While it is more often called the Deming Cycle, Deming referred to this process as the Shewhart Cycle because it was constructed based on work conducted by Walter Shewhart (1939) and can be used to find special cause variation detected by a statistical signal. McSween (2003) specifically applies this continuous improvement cycle to safety culture. First, current safety culture is assessed through observation, so as to fully understand the need for improvement and communicate this need to employees (“Plan”). The opinions and expertise of all levels of management and employees should be incorporated into a team-based implementation plan, and observations of safety practices should occur frequently (“Do”). Corrective feedback should be given and areas for safety improvement should be highlighted (“Check” or “Study”). To promote needed safety improvements and correct behaviors, informational sessions and training should be used appropriately (“Act”). DuPont, a leading company in industrial innovation, has seen great success following this continuous improvement formula for safety (McSween, 2003). Similar applications of continuous improvement methods have been proposed by Proudlove et al. (2008) for healthcare, Reneirs et al. (2009) for chemical safety, and Runyan (2010) for use in the North Carolina State Health Department. The approach is by no means limited to quality and the manufacturing sector.

2.3.3.2 A continuous improvement safety monitoring system

This section proposes a new safety monitoring system to allow for more timely analysis of incident data. The goal of this system is to prioritize incident types for timely root-cause

investigation, so that the most effective hazard controls can be implemented and the largest possible reductions in safety incidents can be achieved as quickly as possible. This safety monitoring system is based on the methodology of the *2mm* Program, a quality improvement application from the United States automobile industry which has experienced great success since the 1990s (Ceglarek and Shi, 1994).

In the *2mm* Program, variation in vehicle body size was reduced by highlighting the measurement points seeing the highest degree of variation and only focusing on reducing the causes of variation at those points (Ceglarek and Shi, 1994). The most significant five percent of problem areas were addressed first. When the number of problems associated with those variation points could be categorized with the other 95%, the most significant five percent was analyzed again and a new set of problem areas were addressed. This continued until the total amount of average body variation was reduced to 2mm, allowing American car manufacturers to remain competitive with their Japanese counterparts (Polenske and Rockler, 2004). Similar processes have also been applied to multi-stage systems, such as compliant sheet metal assembly (Camelio et al., 2003). Shi and Zhou (2009) provide a review of quality control and improvement methodologies for multistage systems.

The following proposed monitoring system combines the themes of the Safety Decision Hierarchy/Hierarchy of Controls, the safety PDCA cycle, and the *2mm* Program to encourage implementation of the most effective hazard controls as quickly as possible. The suggested steps in this methodology follow below, and are illustrated in a cyclical representation in Fig. 2.3.

Identify high priority areas of improvement. As discussed previously, current reporting methods do not provide tools to identify critical areas, unless aggregated historical data is studied to determine the major contributors to accidents. However, the use of accident descriptions can provide sufficient information to detect critical areas. One way of effectively communicating this information is to create a word cloud of accident descriptions or types. Word clouds (also called “tag clouds”) are a popular text mining data visualization tool used to depict the commonality of words in a group of text. As Sinclair and Cardew-Hall (2007) explain, the most abundant words will be represented by larger and bolder text in the cloud, making the highest priority areas of improvement obvious. Common and irrelevant words (“the,” “and,” “employee,” etc.) can be ignored by the word

cloud, and we will designate the highest priority words as the five most predominant words (the largest and boldest in the cloud). While tag clouds have been suggested for data analysis, Sinclair and Cardew-Hall (2007) suggested that they are best used for data portrayal and visualization purposes. This will be their function in this methodology.

Calculate the Safety Incident Indicator value. A key element of the *2mm* program methodology was to define a stable indicator capable of tracking process improvements. In the *2mm* Program, the total number of problems associated with the five percent of high priority variation points was called the Continuous Improvement Indicator. In the case of safety performance monitoring, the indicator can relate to the frequency or severity of incidents. In this paper, the total number of incidents that are represented by the five high priority incident types is called the Safety Incident Indicator value. This is a quantitative indicator, which can be tracked over time to measure the effectiveness of the continuous improvement approach.

Implement hazard controls. Once the most critical areas are identified, it is the responsibility of safety experts to focus efforts on removing the incident root causes. Following the Hierarchy of Hazard Controls, safety professionals should implement the most effective hazard controls possible to reduce the number of incidents specifically associated with the five high priority incident types.

Monitor the Safety Incident Indicator value. How the Safety Incident Indicator value changes over time will indicate the effectiveness of the implemented controls. It is expected that the Indicator value will decrease along with the total number of incidents as more controls are implemented.

Repeat. This process should be repeated regularly in order to always be addressing the highest priority incident types. As the Safety Incident Indicator values decrease and more hazard controls are in place, eventually all incident types will be addressed in order of decreasing priority. Since some companies may transition to this system from annual aggregated reporting methods, it is suggested that the Safety Incident Indicator be updated each month. The time window used to calculate the Safety Incident Indicator must be long enough to capture representative data and produce a stable indicator value, but short enough

to avoid a long delay between the occurrence of an incident and any associated improvement opportunities. If possible, it may be desirable to update the Safety Incident Indicator value more frequently, such as bi-weekly or weekly.

This proposed methodology is depicted in Fig. 2.3; due to the necessary continuous repetition, it is represented as a cyclic process. Section 2.4 presents the expected benefits when applied to an example.

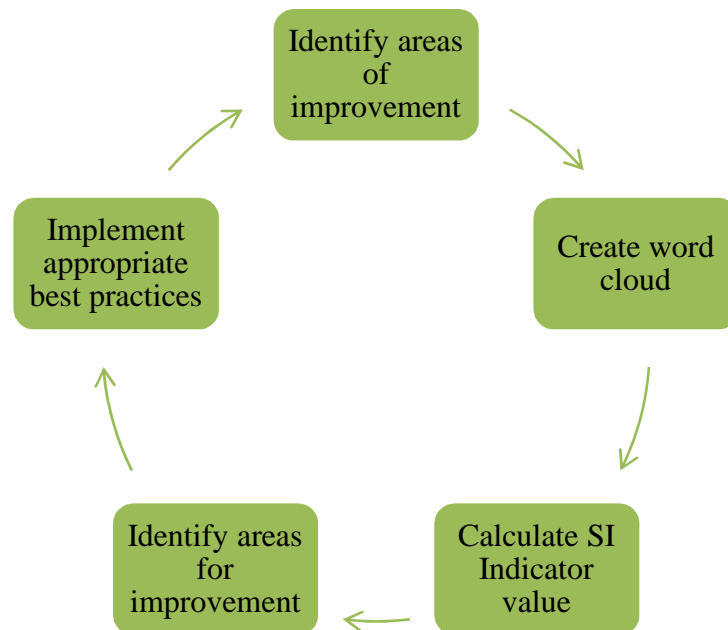


Figure 2.3: Proposed cycle for implementation of the continuous improvement methodology

2.4 Results and Discussion

2.4.1 Validation of the proposed monitoring system

A large historical incident database was provided by an industrial partner and a subset of data was selected. The frequency of incidents remained relatively stable over time, suggesting that there were little or no sources of variation present in the subset. Incident frequency was not increasing, but it was not decreasing either.

In an effort to improve data collection and monitoring, the methodology proposed in Section 3.2 was applied to this example. The high priority incident types from the first month were determined using a word cloud, shown in Fig. 2.4. The most prevalent categories during this

month were identified as “Body Motions,” “Striking,” “Dust, dirt, chips,” “Falling object,” and “Slipping, Tripping, Stumbling.” If available, free-form narrative incident descriptions should be used so that adjectives and other information can be included (Sorock et al., 1997); this would be an example of utilizing the word clouds for data analysis purposes. This type of data was not available for the example data set.

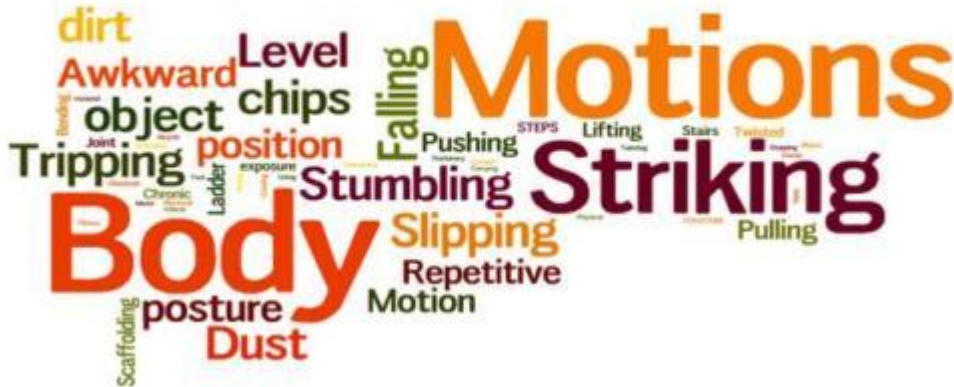
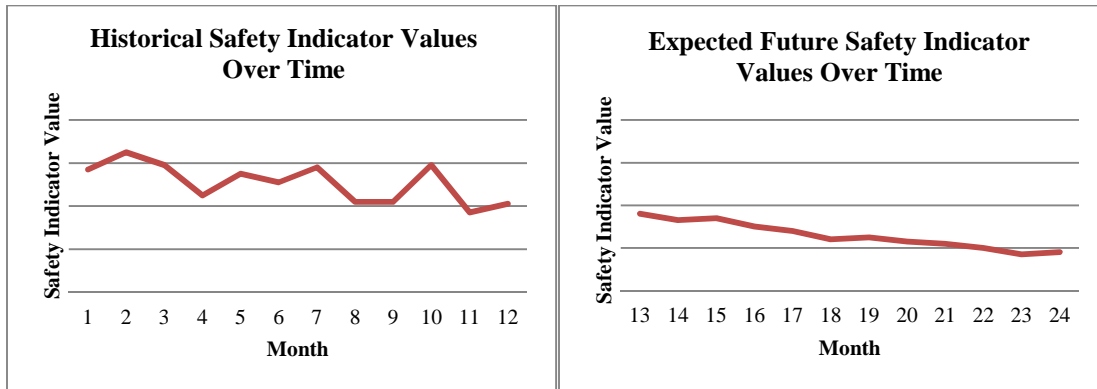


Figure 2.4: Word cloud data visualization of incident types prevalent in industrial data

The Safety Incident Indicator value, the total number of incidents represented by these five accident types for this particular month, was calculated. If this information was made available in a timely manner, it is assumed that root-cause investigation could be conducted for the high priority incident types and appropriate hazard controls could be implemented; these new controls should then cause the Safety Incident Indicator value to decrease. Fig. 2.5a displays the historical monthly values of the Safety Incident Indicator for the dataset; it remains relatively stable over time. As mentioned before, the stability of the performance indicator is critical to track the effectiveness of best practice implementation. Fig. 2.5b shows the expected future pattern of decreasing Safety Incident Indicator values with the proposed active monitoring process in place. A gradual decrease in Indicator values is expected (instead of a sharp decrease) for a few reasons: first, when monitoring over short time periods, the full effects of the implemented controls may not be evident immediately; also, even if the highest priority incident types are effectively controlled, the incident types with the next highest priorities will take their place in the incident count. The actual indicator values have been removed from the y-axes to protect confidentiality.



(a)

(b)

Figure 2.5: Historical Safety Incident Indicator values from Firm X with current performance measurements (a) and expected future Safety Incident Indicator values using the proposed monitoring process (b)

Table 2.2 maps the proposed Safety Incident Indicator active monitoring methodology to the Hierarchy of Hazard Controls. Since it requires timely root-cause investigation of incidents, this proposed methodology will allow safety officials to more practically implement elimination and substitution hazard controls. When elimination and substitution are not possible, the additional information provided by the proposed monitoring system will also allow for more effective engineering controls, administrative controls, and personal protective equipment requirements.

	PPE	Administrative Controls	Engineering Controls	Substitution	Elimination
Proposed active monitoring process	X	X	X	X	X

Table 2.2: Mapping of the proposed active monitoring methodology to perceived levels of achievement on the Hierarchy of Hazard Controls

2.4.2 Near Misses

This new methodology will also allow for the inclusion of “near miss” or “near hit” data; that is, information about incidents that do not lead to a recordable injury could also be considered in the Safety Incident Indicator methodology. Tarrants (1980) notes: “Because most safety measures are postmortem or after-the-fact in nature, they have little but historical value.” If the causes of

near misses can be investigated and addressed before a recordable injury occurs, this will be an obvious step in continuous improvement for safety. Tarrant further explains, "...measures of safety performance must help *prevent* – not *record* – accidents. They must be directional in time and space." Many companies are beginning to collect data on near misses, but enforcing reports of these incidents is often difficult. Wu et al. (2010) specifically encourages real-time near-miss reporting on construction sites.

Most of the currently used performance metrics would not allow for this information to be so easily included. For example, the accident rate does not account for near misses; a site with zero accidents (but many near misses) could be just as 'unsafe' as one with many accidents, but this would not be reflected by their respective accident rates (Mohamed, 2003). Because the Safety Incident Indicator approach suggests targeting areas for improvement based on the descriptions of the most frequent types of injuries, the impact of using leading near-miss indicators to fuel improvements before severe accidents occur will make the approach even more valuable and proactive. Actively monitoring word clouds and Safety Incident Indicator values that include near miss data will make it possible for safety officials to isolate areas needing improvement, even before any accidents occur.

2.5 Conclusions and Future Work

Improving safety on construction sites is an important and relevant topic. Due to the aggregated and lagging nature of current performance measurements, they do not encourage the implementation of the most effective hazard controls. In addition to the obvious concerns for employee well-being, there are cost implications associated with current reactive and lagging systems. Lost time, low EMR, OSHA violations and the like will hurt the bottom line as well. A new continuous improvement active monitoring process is suggested to address the shortcomings in OSHA/BLS reports, Accident and Incident rates, and EMR scores. The proposed process draws upon concepts from the Hierarchy of Hazard Controls, the safety Plan-Do-Check-Act cycle and the automobile industry's continuous quality improvement program. Validating the expected performance of the process with actual data from Firm X suggests that this tool is effective at monitoring safety data in a timely manner, identifying high priority areas of improvement, and assessing hazard controls as they are implemented. Furthermore, the simplicity of including proactive near miss data is explained. For regulatory (e.g. OSHA) and

scientific (e.g. NIOSH) organizations, the system will promote more strategic decision making and resource allocation. When implemented in individual firms, the Safety Incident Indicator methodology has the potential to improve safety environments and reduce the number of injuries and deaths much more effectively than current performance measurements.

Future work might include adapting the process for individual firms, other industries, and other measurements besides frequency (for example, incident severity).

2.6 References

Barnett, R.L., Brickman, D., 1985. Safety hierarchy. Triodyne Safety Systems, L.L.C. Safety Brief, ISSN 1041-9489.

Beavers, J.E., Moore, J.R., Rinehart, R., Schriver, W.R., 2006. Crane-related fatalities in the construction industry. *Journal of Construction Engineering and Management* 132(9), 901-910.

Bureau of Labor Statistics (BLS) (2013). Injuries, illnesses and fatalities. Retrieved from <http://bls.gov/iif/>

Camelio, J., Hu, S. J., Ceglarek, D., 2003. Modeling variation propagation of multi- station assembly systems with compliant parts. *Journal of Mechanical Design* 125, 673-681.

Ceglarek, D., Shi, J., 1994. Dimensional variation reduction for automotive body assembly. *Manufacturing Review* 8(2), 139-154.

Centers for Disease Control and Prevention (CDC). (2012). www.cdc.gov

CPWR (2008). Construction chart book: the US construction industry and its workers. Silver Spring. Retrieved from www.cpwr.com

Deming, W. E. (1986). *Out of the Crisis*. The MIT Press.

El-Mashaleh, M.S., Rababeh, S.M., Hyari, K.H., 2010. Utilizing data envelopment analysis to benchmark safety performance of construction contractors. *International Journal of Project Management* 28, 61-67.

Everett, J.G., Thompson, W.S., 1995. Experience modification rating for workers' compensation insurance. *Journal of Construction Engineering and Management* 121(1), 66-79.

Hinze, J., Pedersen, C., Fredley, J., 1998. Identifying root causes of construction injuries. *Journal of Construction Engineering and Management* 124(1), 67-72.

- Hinze, J., Russell, D.B., 1995. Analysis of fatalities recorded by OSHA. *Journal of Construction Engineering and Management* 121(2), 209-215.
- Hinze, J., Thurman, S., Wehle, A., 2013. Leading indicators of safety performance. *Safety Science* 51, 23-28.
- Hoonakker, P., Loushine, T., Carayon, P., Kallman, J., Kapp, A., Smith, M.J., 2005. The effect of safety initiatives on safety performance: A longitudinal study. *Applied Ergonomics* 36, 461-469.
- Huang, X., Hinze, J., 2003. Analysis of construction worker fall accidents. *Journal of Construction Engineering and Management* 129(3), 262-271.
- Manuele, F.A., 2005. Risk assessment & hierarchies of control. *Professional Safety* 50(5), 33-39.
- McSween, T.E., 2003. *Values-based safety process: improving your safety culture with behavior-based safety*, 2nd Edition. Hoboken, NJ: John Wiley & Sons.
- Mohamed, S., 2003. Scorecard approach to benchmarking organizational safety culture in construction. *Journal of Construction Engineering and Management* 129(1), 80-88.
- Ng, S., Cheng, K., Skitmore, R., 2005. A framework for evaluating the safety performance of construction contractors. *Building and Environment* 40(10), 1347-1355.
- Occupational Safety and Health Administration (OSHA) (2013). "Commonly Used Statistics." Retrieved from <http://www.osha.gov/oshstats/commonstats.html>
- Polenske, K.R., Rockler, N.O., 2004. Closing the competitive gap: A retrospective analysis of the ATP 2mm Project. NIST publication GCR 03-856.
- Proudlove, N., Moxham C., Boaden, R., 2008. Lessons for Lean in healthcare from using Six Sigma in NHS. *Public Money and Management* 28(1), 27-34.
- Reniers, G.L.L., Ale, B.J.M., Dullaert, W., Soudan, K., 2009. Designing continuous safety improvement within chemical industrial areas. *Safety Science* 47, 578-590.
- Runyan, C.W., 2010. Positioning North Carolina for leadership in injury control: A call to action." *North Carolina Medical Journal* 71(6), 527-530.
- Shewhart, W.A., 1986. *Statistical method from the viewpoint of quality control*. Reprint from original publication: Washington D.C., Graduate School of the Department of Agriculture, 1939.
- Shi, J., Zhou, S., 2009. Quality control and improvement for multistage systems. *IIE Transactions* 41, 744-753.

Sinclair, J., Cardew-Hall, M., 2007. The folksonomy tag cloud: when is it useful? *Journal of Information Science* 34, 15-29.

Sorock, G.S., Smith, G.S., Reeve, G.R., Dement, J., Stout, N., Layne, L., Pastula, S.L., 1997. Three perspectives on work-related injury surveillance systems. *American Journal of Industrial Medicine* 32, 116-128.

Stevenson, W.J., 2009. *Operations Management*, 10th Edition (p. 428). New York: McGraw Hill.

Tam, C.M., Fung, W.H.I., 1998. Effectiveness of safety management strategies on safety performance in Hong Kong. *Construction Management and Economics* 16, 49-55.

Tarrants, W. E., 1980. *The measurement of safety performance*. New York: Garland STPM Press.

U.S. Department of Defense, 2012. MIL-STD-822E Standard practice for system safety.

Wu, W., Yang, H., Chew, D.A.S., Yang, S., Gibb, A.G.F., Li, Q., 2010. Towards an autonomous real-time tracking system of near-miss accidents on construction sites. *Automation in Construction* 19, 134-141.

This is an Author's Accepted Manuscript of an article by Anna Schuh, Jaime A. Camelio, and William H. Woodall, which has been accepted for publication in the *International Journal of Injury Control and Safety Promotion*, published by Taylor & Francis. The *International Journal of Injury Control and Safety Promotion* is available online at <http://www.tandfonline.com/loi/nics20>.

3. Control Charts for Accident Frequency: A Motivation for Real-Time Occupational Safety Monitoring

3.1 Abstract

This paper quantitatively motivates the need for active monitoring of occupational safety incident data through the use of cumulative sum (CUSUM) control charts. The frequency of incidents within a subset of historical accident data is analyzed. The performance of Poisson CUSUM and exponential CUSUM (time-between-events) charts is compared in an illustrative example to show that shorter periods of aggregation and time-between-events monitoring lead to more timely indications of increased accident frequency. An extension showing the anticipated performance of these charts with real-time data is given. Various adjustments to the monitoring system are also simulated to show that quick implementation of hazard controls can significantly impact safety performance. Decreases in the frequency of safety incidents as a result of implemented hazard controls can also be monitored.

3.2 Introduction

An important issue in occupational safety is the timely detection of changes in safety incident patterns. In 2011, 3.5 per 100,000 employees were killed on the job in the United States, and the average non-fatal injury resulted in 8 days away from work (Occupational Safety and Health Administration, OSHA, 2013). If increases in accident frequency are noted and new hazard controls are implemented quickly, these efforts could lead to fewer employee injuries and deaths in the workplace. The National Institute for Occupational Safety and Health (NIOSH) defines occupational health surveillance as “the tracking of occupational injuries, illnesses, hazards, and exposures” used to “guide efforts to improve worker safety and health, and to monitor trends and progress over time” (NIOSH, 2011). The Institute has developed several surveillance programs, such as the Fatality Assessment and Control Evaluation (FACE) Program for monitoring fatal incidents and the Occupational Respiratory Disease Surveillance system. Achieving better

surveillance practices is one of the strategic goals of the National Occupational Research Agenda (NORA, 2012).

In an effort to provide better surveillance of occupational safety incidents, large amounts of data are collected by various national organizations such as the Occupational Safety and Health Administration (OSHA) and the Bureau of Labor Statistics (BLS). However, companies are only required to submit information to OSHA about injury-causing incidents aggregated over a period of one year, and BLS reports this aggregated data every September for the previous year (OSHA, 2013; BLS, 2013). Individual companies will experience the effects of the time lag between their annual OSHA reports, unless worker compensation programs or company-directed safety initiatives require more frequent incident monitoring or analysis. Root-cause investigations occur at the company or plant level.

This paper demonstrates how real-time accident monitoring, instead of aggregated reporting, can more quickly indicate increases in accident rates and encourage the efficient implementation of appropriate best practices and hazard controls. Time-between-events (TBE) cumulative sum (CUSUM) control charts are used to monitor changes in the frequency of incidents by using the times between successive incidents. The performance of these charts is compared to that of Poisson CUSUM control charts for aggregated incident counts of various time lengths to show that real-time or near-real-time incident monitoring detects increases in accident frequency much more efficiently than long-term aggregated reporting.

In industrial process control, a common tool for monitoring defective parts in manufactured products is the control chart (Montgomery, 2008). Control charts have been extended to safety applications before, but not for the purpose of motivating real-time occupational incident monitoring. For example, control charts have been used to investigate trends in traffic accidents: Blindauer and Michael (1959) used traditional Shewhart control charts to investigate changes in accident rates on certain highway sections in Indiana; Guria and Mara (2000) suggested Shewhart charts for monitoring weekly and monthly traffic accidents in New Zealand; Sparks (2010) introduced a two-pass CUSUM chart to monitor monthly traffic crash outbreaks for various age groups in New Zealand; and Adekeye and Aluko (2012) used a CUSUM chart to identify the time periods prone to the highest rate of road accidents in Nigeria. Similarly, the use of control charts for monitoring chemical accidents has been discussed: Cournoyer et al. (2011)

used a Shewhart c-chart to show increases in monthly chemical injuries. Benneyan (1998a, 1998b), Woodall (2006), and Woodall et al. (2010) gave extensive reviews about how control charting can also be used to monitor a wide variety of attributes in healthcare. TBE CUSUM monitoring for the frequency of industrial accidents was specifically suggested by Lucas (1985), but the time-between-incidents rate was estimated based on monthly aggregated accident totals. Wu et al. (2010) suggested a real-time TBE chart to monitor occupational fatalities, but did not consider non-fatal injuries or other safety incidents.

In this paper, the use of control charts is extended to real-time monitoring of the frequency of all occupational safety incidents, to encourage the quick recognition and implementation of new hazard controls. It is assumed that these efforts will lead to a subsequent reduction in the number of occupational injuries.

The paper is organized as follows. First, past accidents from an industrial safety incident database are considered with Poisson cumulative sum (CUSUM) charts for various levels of aggregation (quarterly, monthly, semi-monthly, weekly, every three days, and daily). The time to signal of these charts is compared to that of a time-between-events (TBE) exponential CUSUM chart in Section 3.3. In Section 3.4, an example of the expected performance of the TBE and daily Poisson CUSUM charts when used to actively mitigate hazards with real-time data is simulated. The performance of different variations of the monitoring system is reported, including ineffective hazard controls, shorter periods for control implementation, and nonzero starting and restarting values for the CUSUM charts. Finally, a discussion of current limitations, future work, and conclusions are presented.

3.3 Cumulative Sum Control Charts for Monitoring Incident Data

In this section we introduce the background and structure of the Poisson and TBE CUSUM charts. Then, in Section 3.4, a dataset based on industrial accidents will be used to construct Poisson CUSUM charts with incident counts reported quarterly, monthly, semi-monthly, weekly, every-three-days, and daily. The time between incidents will also be used to construct an exponential CUSUM chart.

The dataset used in this paper was provided by a large industrial partner. A smaller subset of the data, those incidents associated with one particular injury type, has been selected to protect

confidentiality. The dataset has a total of 23 incident categories, but since the frequency of incidents is of current interest, the accident date is the only data field used for this analysis (along with the injury type filter used to define the subset). The number of days between incidents and the total number of incidents in each of the proposed aggregation periods had to be calculated from the original data. This was the only data pre-processing necessary for this work.

3.3.1 Poisson CUSUM charts

Traditional control charts use a baseline dataset of typical data (Phase I) to establish charting parameters, and then actively monitor new data points as they are observed (Phase II). When the defect count surpasses an established upper limit, the chart is said to “signal” that the process is “out-of-control.” The process is stopped until the source of the problem is identified and corrected. The first control charts were used to monitor count data, usually the number of nonconforming parts, and signal out-of-control conditions. These charts were proposed by Walter A. Shewhart and are therefore referred to as Shewhart charts.

The CUSUM control chart, first proposed by Page (1954), incorporates more of the information in the data. Because historical data is included, CUSUM charts are more sensitive to small shifts in underlying parameters and better at identifying the point of change than standard Shewhart c-charts (White et al., 1997).

The values of counted data, such as the number of nonconforming parts or the number of accidents, often follow a Poisson distribution (Lucas, 1985). Therefore, a Poisson CUSUM chart is appropriate for monitoring counted data. To implement a Poisson CUSUM chart, the total number of adverse events is aggregated over a fixed time period. The CUSUM statistic at time i , $S_{\text{POISSON},i}$, used to detect an increase in the rate of adverse events, is given by

$$S_{\text{POISSON},i} = \max(0, S_{\text{POISSON},i-1} + X_i - k_{\text{POISSON}}), \quad (1)$$

where X_i is the observed number of events during time period i and k_{POISSON} is the reference parameter based on the Phase I data. The value of $S_{\text{POISSON},0}$ is usually set to zero. If μ_a is the historical mean rate (usually chosen to be equal to or near the mean of a baseline “Phase I” dataset of data) and μ_d is the mean rate to be detected quickly, then the reference value for the Poisson CUSUM can be found by

$$k_{POISSON} = \frac{\tilde{d} - \tilde{a}}{\ln(\tilde{d}) - \ln(\tilde{a})} \quad (2)$$

The chart signals when $S_{POISSON,i} > h_{POISSON}$. The decision interval, $h_{POISSON}$, can be found using table look-up procedures (Lucas, 1985; Osanaiy and Talabi, 1989), computer programs (Hawkins and Olwell, 1998), or using a general rule-of-thumb such as four-to-five times the process standard deviation (Montgomery, 2008). The Poisson CUSUM chart has been applied in non-manufacturing settings, such as monitoring automobile accidents (Adekeye and Aluko, 2012) and diabetic epidemics (Osanaiy and Talabi, 1989).

3.3.2 Time-Between-Events Control Chart

An alternative to monitoring aggregated incident counts is to study the amount of time between events, or the inter-arrival time. The exponential CUSUM control chart (also known as a “time-between-events” or “time-between-occurrence” chart) has been proposed for this purpose. The amount of time between Poisson-distributed events follows an exponential distribution for a Poisson process, and the time-between-events (TBE) chart monitors these times (Gan, 1994). TBE monitoring is based on the amount of time that has lapsed between the current event and the event immediately prior. Thus, rather than waiting until the end of a fixed time period for aggregated count data, the information is incorporated as it is obtained during the process. When monitoring the occurrence of rare adverse events (such as nonconforming parts, surgical errors, or industrial accidents) an increase in the time between events is desirable and indicates an improvement in the process. Decreases in the time between events should be detected and investigated as quickly as possible. Lucas (1985) recommended the TBE method for monitoring industrial accidents, if it is plausible for practitioners to record events as they happen instead of aggregating counts. In a limited study without much aggregating over time, Gan (1994) showed that zero state Poisson CUSUM and exponential CUSUM charts have similar performance for small decreases in the average time between events, but the exponential CUSUM is better for detecting large decreases. Schuh et al. (2013) have extended Gan’s (1994) work by comparing the steady state ATS performance of charts with even longer aggregation periods than Gan (1994) considered.

The CUSUM statistic for the TBE chart at time i is defined as

$$C_{TBE,i} = \max (0, C_{TBE,i-1} + k_{TBE} - X_i) \quad (3)$$

where $C_{TBE,0} = 0$ and the reference value is given by

$$k_{TBE} = \frac{\lambda_d \lambda_a}{\lambda_d - \lambda_a} \ln\left(\frac{\lambda_d}{\lambda_a}\right) \quad (4)$$

where λ_a is the average number of time periods between accidents during Phase I, and λ_d is the average to be detected (Gan, 1994). Since a decrease in the expected time periods between accidents is to be detected, $\lambda_d < \lambda_a$. Lucas (1985) used monthly rates with his TBE CUSUM chart, so he instead considered the Poisson count of accidents per time period (μ_a and μ_d , the reciprocals of λ_a and λ_d). This k_{TBE} is calculated equivalently by

$$k_{TBE} = \frac{\ln(\lambda_d) - \ln(\lambda_a)}{\lambda_d - \lambda_a}.$$

The chart signals when $C_{TBE,i} > h_{TBE}$ and the decision limit h_{TBE} can again be determined by a number of methods.

The TBE charts, like all control charts, have traditionally been used in manufacturing, such as Khoo and Xie's (2009) application to observed failures in regularly maintained systems. As mentioned, Lucas (1985) was the first to discuss the use of TBE CUSUM charts for industrial accident monitoring using event rates based on monthly aggregated incident totals; Wu et al. (2010) used hours between events for a real-time TBE CUSUM chart, but only included accident fatalities. Khanzode et al. (2011) suggested TBE charts for identifying accident trends and mitigating hazards in coal mining.

In Section 3.4, the performance of Poisson and TBE CUSUM control charts will be compared for one example of occupational accident data. This expands upon the work of Lucas (1985) by considering the time-between-events, rather than assumed rates based on monthly aggregated totals, and complements the work of Wu et al. (2010) by considering all accidents rather than only fatalities. The effects of several variables are also considered, including the relative success of implemented hazard controls, shorter control implementation periods, and non-zero starting and re-starting CUSUM values.

3.4 Application to Industrial Accident Data

In this section we consider the effects of data aggregation for the provided Poisson dataset. The effects of aggregated data have been considered before, but not for occupational safety applications and mainly in the case of comparing aggregated binomial samples against real-time Bernoulli data. In their review on the surveillance of Bernoulli processes, Szarka and Woodall (2011) discussed the possibility of chart inefficiency when aggregated binomial data is used instead of Bernoulli data. Reynolds and Stoumbos (1999, 2000, 2004) considered the problem of finding the best sample size when aggregating Bernoulli outcomes related to parts in manufacturing and determined that for CUSUM charts smaller sample sizes provide the best chart performance, signaling trends sooner than with the use of large sample sizes. Since discrete Bernoulli data with geometrically distributed wait times is analogous to the continuous exponential waiting times scenario considered here, the conclusions should be similar.

Data aggregation has been briefly addressed for monitoring in healthcare. Sego et al. (2008) noted for the surveillance of birth defects that Poisson or binomial data aggregation can reduce the efficiency of control charts. For other clinical applications, Benneyan (2008) suggested collecting data in frequent, smaller time periods.

First, the performance of the Poisson and TBE CUSUM charts is compared using historical data.

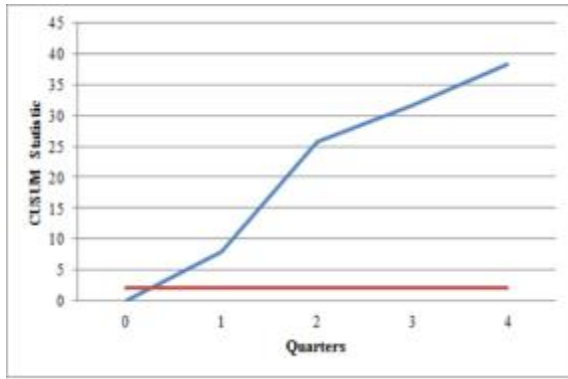
3.4.1 Historical Data

Using the dataset provided by the industrial partner, the time-to-signal was analyzed for Poisson CUSUM charts with quarterly, monthly, semi-monthly, weekly, and daily counts of injuries, as shown in Figure 3.1. This example is based on the assumption that the data would be available in a continuous stream and able to be counted and reported regularly in any of these time periods. For each monitoring period, one year of “typical” data was used as the Phase I baseline to calculate the reference value k_{POISSON} using Eqn. 2. Year 2 data was then monitored for the various levels of aggregation in real time. The mean number of incidents to be detected, μ_d , is assumed to be one standard deviation above the Year 1 average, μ_a , for each of the aggregation periods considered. For instance, in the monthly chart, the average monthly rate of accidents in Phase I was 4.5 accidents per month (μ_a). The rate to be detected, μ_d , was one standard deviation higher than μ_a , 6.7 accidents per month. The decision interval, h_{POISSON} , was determined in each

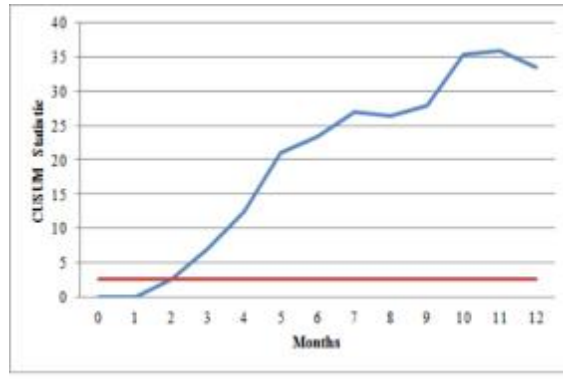
case using the tables provided by Lucas (1985) to give a high average run length (ARL) value of approximately one year when the mean is in-control and low ARL values when out-of-control. As Lucas (1985) points out, the ARL of a Poisson chart is the average number of time intervals before an out-of-control signal is given; so one year for the monthly chart would mean an ARL of 12 periods, or 24 periods for the semi-monthly chart, etc.

When interpreting Figure 3.1, the quarterly aggregation produces a signal after the first quarter (month 3, after the 23rd accident); monthly after month 3 (23 incidents); semi-monthly during the middle of month 3 (11 weeks, 19 accidents); weekly after 10 weeks (16 accidents); every 3 days after 10 weeks (16 incidents); and daily during the tenth week (15th accident). In a monitoring system with real-time data, the CUSUM statistic should be reset following a signal and corrective actions taken; this step is not shown in Figure 1 since only the time-to-signal is of interest, but resetting will be considered in the next subsection.

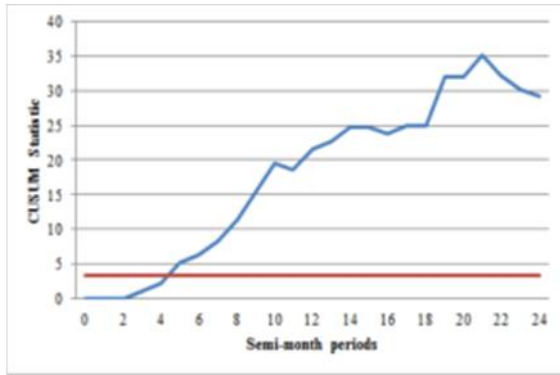
When using aggregated Poisson count data, it is important to carefully choose the length of the monitoring period. Obviously, longer monitoring periods lead to a higher degree of aggregation and a subsequent loss of information. If practitioners need to wait until the end of a longer time period (a quarter, or a month, etc.) to detect an increase in injuries, the time to effectively identify and control associated hazards may have passed. Furthermore, new and as-yet-unrecognized hazards may have developed that are not mitigated by previously implemented controls, and this dynamic may not be captured by the aggregated count. Our results indicate that charts using data from shorter monitoring periods detect increases in accident frequency more quickly. At least in our example, the daily Poisson CUSUM chart signals an increase in the number of safety incidents sooner than any of the charts with longer monitoring periods, although with a case study example we cannot know if this is a false alarm or a true indication of an increased accident rate.



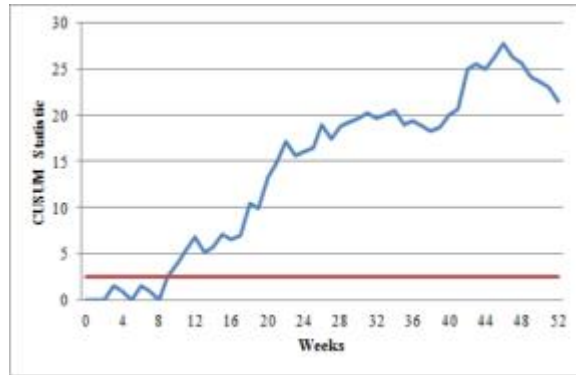
(a)



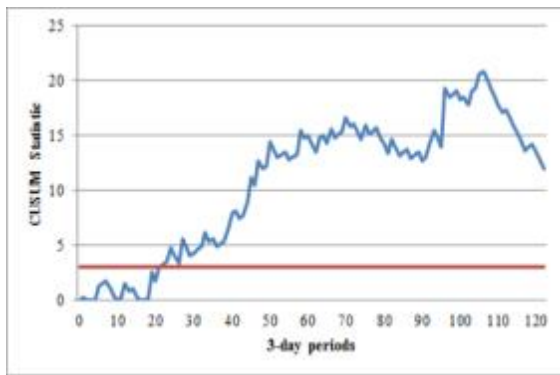
(b)



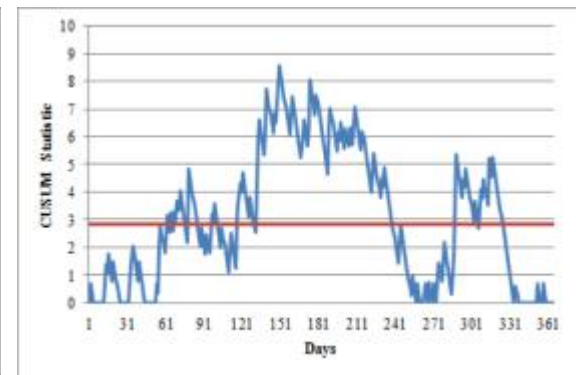
(c)



(d)



(e)



(f)

Figure 3.1: Year 2 Poisson CUSUM control charts for aggregated accident counts, a) quarterly; b) monthly; c) semi-monthly; d) weekly; e) every 3 days; f) daily. The x-axis labels represent time as the number of specified time periods passed. The y-axis is the cumulative sum value.

In addition to a providing the most timely signal, fluctuations are evident with the daily Poisson CUSUM chart that are not obvious with other monitoring periods. The decrease observed around

day 241 of Figure 3.1(f) may have been due to an implemented hazard control or best practice; the apparent positive effect of this measure would not be noticed with longer period lengths.

The same dataset was used to construct a TBE control chart, shown in Figure 3.2, where the number of days between subsequent accidents is considered. On days when multiple accidents are reported, the time between the first and second accidents (and second and third, etc.) is recorded as “0 days.” The time to signal for this chart is the same as the daily Poisson CUSUM chart (during the tenth week, after 15 accidents), as shown in Figure 3.2. Aggregation by day is likely the shortest reasonable level of aggregation for most companies. This affirms that the aforementioned findings of better performance with shorter periods of data aggregation and real-time time-between-accidents (Reynolds & Stoumbos, 1999, 2000, 2004; Benneyan, 2008) are also applicable for this example of a safety-related Poisson process. The same one year of baseline Phase I data was used and reference parameter, k_{TBE} , was calculated using Eqn. 4. The in-control time between events, μ_a , was the average Year 1 time between events of 6.5 days between events. The time-between-events value to be detected, μ_d , was calculated to be one standard-deviation less than μ_a (3.4 days between events). The decision interval, h_{TBE} , was determined using the tables provided by Lucas (1985).



Figure 3.2: Year 2 time-between-events CUSUM control chart

The TBE CUSUM communicates little information past the signal after the 15th accident, since the average number of accidents continues to be higher than Phase I. However, when used for

prospective safety surveillance, a control signal should indicate the immediate need for hazard investigation. The sooner hazards are identified, the more likely it is that the hazard control programs can have the largest effect. After the identified problem is resolved, the CUSUM monitoring statistic will be reset to its initial value. The illustrative performance of both the daily Poisson and TBE CUSUM charts with our real-time data is given in the following section.

3.4.2 Discussion of real-time data simulation results

In this section we present simulated daily Poisson and TBE CUSUM charts, representing how the chart for Year 2 data would be expected to look if hazards are actively mitigated as out-of-control conditions are signaled and identified. The simulation is based on the assumption that there is a team of safety officers in charge of monitoring and actively mitigating hazards during Year 2. Then Figures 3.3a and 3.3b show the performance of the daily Poisson and TBE charts respectively, using the same Year 2 data used to construct Figures 3.1 and 3.2. Following the original signal after the fifteenth accident, it is posited that the active monitoring system would motivate the implementation of hazard controls. In manufacturing systems it is assumed that the process would be stopped following a control chart signal, the machine recalibrated or other action taken, the process resumed, and the CUSUM monitoring statistic reset to zero. However, in safety applications, it is not possible to “stop the process” of employees working while a new hazard control is implemented. Therefore, in the simulation, monitoring continues and it is assumed that it will take thirty days following a signal to identify the hazard(s) and implement new best practices. After this thirty day control implementation period, the CUSUM statistic is reset to zero.

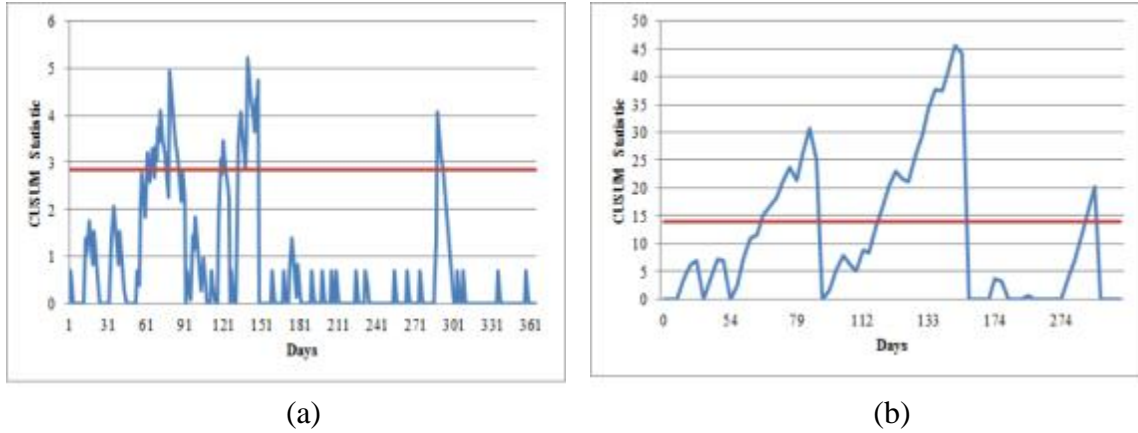


Figure 3.3: Simulated future performance of a) the daily Poisson CUSUM chart and b) the time-between-events CUSUM chart with a thirty day period for control implementation

3.4.2.1 Successful and unsuccessful hazard controls

Monitoring in real-time will provide a quantitative assessment of a new best practice's success or failure. In Figures 3.3a and 3.3b, it is assumed that hazard controls would be implemented in an effort to mitigate the risks associated with increase in accident frequency signaled after day 62 (accident 15); however, it is also assumed that these hazard controls are ineffective and don't prevent any accidents (that is, all incidents from the original database used in Section 3.4.1 are still included here, as if the hazard controls did not reduce the accident frequency at all). The TBE CUSUM signals another increase in incident frequency after day 118 (33 incidents), and the Poisson CUSUM signals after day 119 (accident 34). In both cases, another thirty days is given for hazard control implementation before the CUSUM is reset, and this time it is assumed that 50% of incidents are eliminated with the new controls (so 50% of the incidents used in Section 3.1 are included in this simulation, after the respective second signals for each chart). Third signals are seen after day 288 (accident 65) in both charts, and accidents are assumed to be reduced another 25% by the new controls motivated by these signals. With all of these assumptions of accident reductions, both simulated charts in Figure 3.3 includes a total of 69 accidents, out of a total 99 accidents included in the original Section 3.4.1 dataset (implying a 30% reduction in the total number of accidents using Poisson daily or TBE monitoring).

3.4.2.2 Shorter control implementation periods

In Figures 3.4a and 3.4b, the same simulation is conducted, with only a 10-day control implementation period. This leads to quicker follow-up signals when controls are ineffective;

this time, signals are seen after the 117th (TBE) and 119th (Poisson) days, following the simulated ineffective controls implemented in response to the signal on the 62nd day. A 50% reduction in accident frequency is assumed following these second signals, and another 25% reduction is assumed after the third signals on day 288. For the particular set of data used, no significant improvements were observed; one fewer accident is observed compared to the simulation reported in Section 3.4.1. Of the 99 original accidents, this version of the Poisson CUSUM includes 65 accidents and the TBE CUSUM includes 64 accidents (for a total simulated reduction in accidents of about 33% using the shorter implementation period of 10 days). Use of the TBE CUSUM chart seems to slightly reduce the simulated frequency of incidents over the Poisson daily CUSUM with this modification.

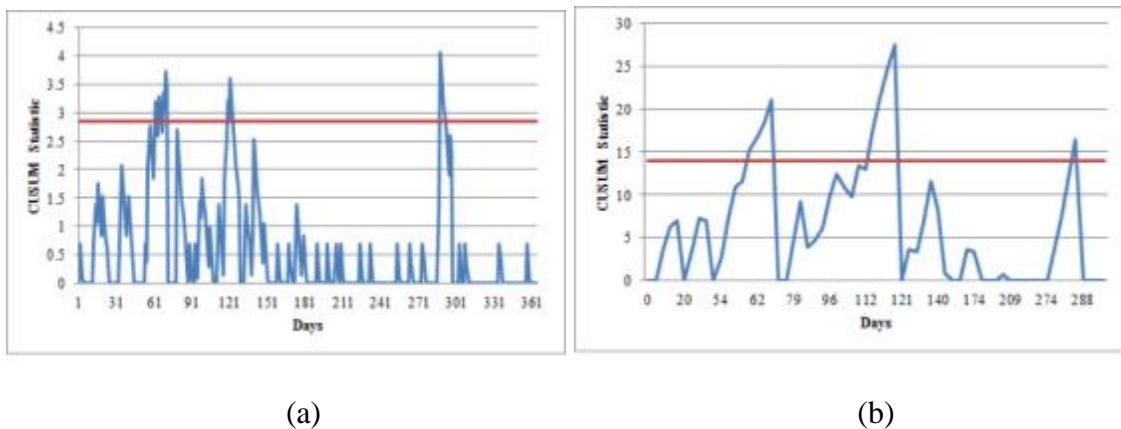


Figure 3.4: Simulated future performance of a) the daily Poisson CUSUM chart and b) the TBE CUSUM chart with a 10-day period for control implementation

3.4.2.3 Non-zero starting and restarting values

Coory et al. (2007) and Grigg et al. (2003) note that in public health it may also be worthwhile to define starting and reset values higher than 0, since it cannot ever be assumed that the process is in control when monitoring begins or restarts. Using the same 10-day control implementation period, Figure 3.5 shows the performance of the daily Poisson and TBE CUSUM charts with starting and reset values of one-half the decision limit. Of the original 99 accidents, 60 accidents for the Poisson chart and 55 for the TBE are now included (total simulated reductions of 39% and 44% with both a 10-day implementation period and non-zero starting and restarting values). Again, there is a greater reduction seen when using the real-time TBE chart. An alternative to using non-zero restarting values is to use an adaptive CUSUM chart (Sparks, 2000) to monitor

steady state data, which will be more likely in public health and safety applications. Many researchers have proposed methods which adapt the design parameters or sampling policies to achieve improved control chart performance for a delayed shift (Tsung and Wang, 2010). Schuh et al. (2013) consider the effects of aggregation when monitoring a Poisson process under a steady state assumption.

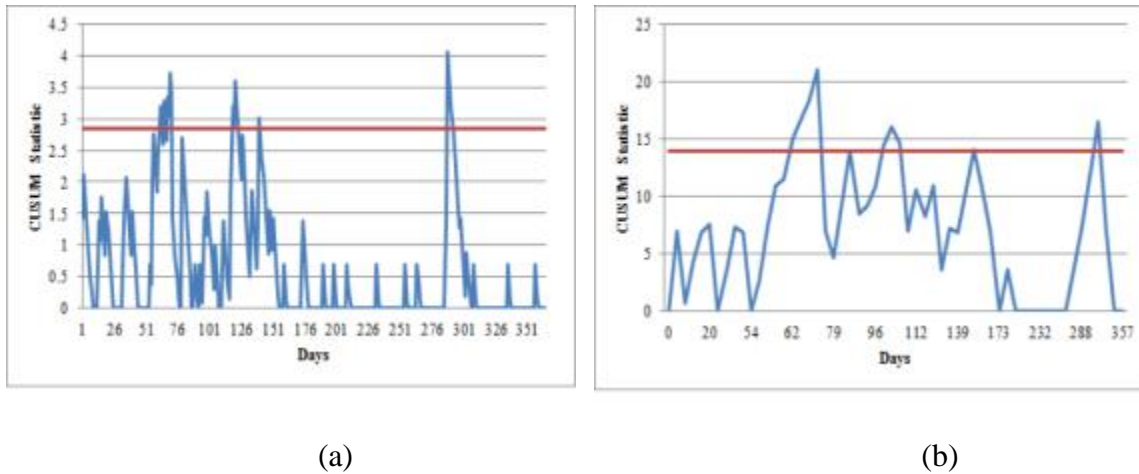


Figure 3.5: Simulated future performance of a) the daily Poisson CUSUM chart and b) the TBE CUSUM chart with adjusted starting and re-starting values

If control charts tracking either the time between incidents or the daily count of incidents are actively monitored by a team of safety officers and each signal is quickly addressed by the implementation of new hazard controls, it is clear how the frequency of accidents could be reduced and, potentially, lives could be saved. It has been shown that an inherent benefit of this system is the ability to determine success or failure of implemented hazards through the subsequent increase or decrease in accident frequency following their implementation. It has also been shown that implementing hazard controls more quickly following a signal (10 days, instead of 30 days) may allow for a more significant reduction in accident frequency.

3.5 Conclusions

This paper has investigated the use of TBE and Poisson CUSUM charts for occupational safety monitoring. It has been shown that TBE control charting provides timely signals of increasing frequency of safety incidents; therefore, we recommend the collection of safety incident data as often as possible for the effective mitigation of hazards. Poisson CUSUM charts were considered for aggregated reporting, and it has been shown that shorter periods of aggregation can lead to

more efficient performance. An example of simulated expected future performance shows that daily Poisson CUSUM and TBE CUSUM charts not only indicate an increase in the incident rate, but also success or failure of implemented hazard controls and best practices. The simulation also shows that quick implementation of hazard controls following a signal may lead to a greater reduction in the number of future accidents.

3.6 Current Limitations and Opportunities for Future Work

While the results clearly encourage real-time safety monitoring, it is necessary to discuss some of the limitations present in the incident database used for this work. Because this work has presented a broad investigation into the use of control chart monitoring for occupational safety, it was necessary to make some assumptions. The specific effects of these assumptions for occupational safety monitoring will need to be investigated at length in future studies. These issues are summarized below.

3.6.1 Data aggregation

One limitation of safety data is its availability; it is assumed throughout this work that the aggregated counts and times between events are readily available as often as needed. In applications, the data may need to be processed; this may not be quick, especially for companies transitioning to a real-time monitoring system from long-term aggregated reporting. New data entry software and trained employees may need to be introduced into the safety system in order to perform the active monitoring described in this work. While this may involve some investment of both time and money, the potential difference in accident rate change detection speed demonstrated by this work should motivate companies to process data as soon as possible.

3.6.2 Data Standardization

Some limitations in the data provided by safety databases have been discussed for road traffic accidents (Peden and Toroyan, 2005; Sethi and Zwi, 1999). A limitation to be considered in occupational safety data is that the data should be standardized; that is, changes in the overall workforce population will effect the frequency of accidents, independent of the level of hazard or safety, and should therefore be accounted for in the analysis. These changes may include the overall population of workers, as well as any changes in the average employee age, level of

training and experience, level of obesity, work seasonality, or other factors which may effect the employees' susceptibility to injury. However, the overall workforce population data may not always accompany an incident database, as is the case with the industrial partner involved in this work. In the example used here, only two back-to-back years' worth of data is used, so it is assumed that the workforce population did not change significantly during this time.

3.6.3 Expected Data Variation

When analyzing safety data, it is expected that there may be inherent day-of-the-week and time-of-the-day trends in incident patterns. Likewise, there may be geographical patterns associated with the work location or the specific area within a large worksite. These common-cause variations should be accounted for within a monitoring system, in order to better detect those unusual patterns with assignable causes.

3.6.4 Incident Severity

Another aspect of safety incidents which should also be considered is their severity. If accident prevention strategies need to be prioritized, emphasis should be placed on preventing those more severe incidents. Langley (2004) posited that a core challenge to injury surveillance is defining and agreeing upon what constitutes an "important injury event," and investigating relative severity may be one way to accomplish this. The Department of Energy has developed a Safety Cost Index, defining the severity of an occupational accident by its associated cost (Gochfield and Mohr, 2007). Wu et al. (2010) considered severity to be the number of fatalities resulting from one catastrophe. Another definition of severity for a non-fatal incident could be the resulting number of days away from work and/or in job restriction.

3.6.5 Near Misses

"Near miss" or "near hit" data can also be included in daily Poisson CUSUM or TBE charts. If incidents which do not lead to recordable injuries are also recorded and monitored with real-time control charting tools, observed increases in the frequency of these incidents may lead to hazard identification and mitigation before significant injuries even occur. In this way, a frequency control chart signal can be a proactive leading indicator. Lander et al. (2011) has shown that

active investigation of near miss claims can lead to an overall decrease in the occupational accident rate. One must be concerned, however, with the under-reporting of near-miss accidents.

3.6.6 More General Results

We present an illustrative example in this paper that shows the benefits of using time-between-events data. Schuh et al. (2013) performed a thorough simulation study to show the effects of data aggregation when monitoring a Poisson process for increases in the rate of adverse events such as accidents.

3.7 References

Adekeye, K.S., & Aluko, O.S. (2012). Design of cusum scheme for monitoring road accident fatalities. *Open Journal of Statistics*, 2, 213-218.

Benneyan, J.C. (1998a). Statistical quality control methods in infection control in hospital epidemiology, Part I: introduction and basic theory. *Infection Control and Hospital Epidemiology*, 19, 194-219.

Benneyan, J.C. (1998b). Statistical quality control methods in infection control in hospital epidemiology, Part II: chart use, statistical properties, and research issues. *Infection Control and Hospital Epidemiology*, 19(4), 265-283.

Benneyan, J.C. (2008). The design, selection, and performance of statistical control charts for healthcare process improvement. *International Journal of Six Sigma and Competitive Advantage*, 4(3), 209-239.

Blindauer, H.H. and Michael, H.L. (1959). An analysis of high accident rates: technical paper. Publication FHWA/IN/JHRP-59/14. Joint Highway Research Project, Indiana Department of Transportation and Purdue University, West Lafayette, IN.

Bureau of Labor Statistics (BLS) (2013). "Injuries, Illnesses and Fatalities." Retrieved from <http://bls.gov/iif/>

Coory, M., Duckett, S., and Sketcher-Baker, K. (2007). Using control charts to monitor quality of hospital care with administrative data. *International Journal for Quality in Health Care*, 20(1), 31-39.

Cournoyer, M., Garcia, V., Gallegos, U., and Wilburn, D. (2011). Investigation of injury/illness data at a nuclear facility. *Journal of Chemical Health and Safety*, 8(5), 17-25.

Gan, F.F. (1994). Design of optimal exponential CUSUM control charts. *Journal of Quality Technology*, 26(2), 109-124.

Gochfield, M. and Mohr, S. (2007). Protecting contract workers: case study of the U.S. Department of Energy's nuclear and chemical waste management. *American Journal of Public Health*, 97(9), 1607-1613.

Grigg, O.A., Farewell, V.T., and Spiegelhalter, D.J. (2003). Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. *Statistical Methods in Medical Research*, 12(2), 147-170.

Guria, J. and Mara, K. (2000). Monitoring performance of road safety programmes in New Zealand. *Accident Analysis and Prevention*, 32, 695-702.

Hawkins, D. M. and Olwell, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*. Springer Verlag, New York, NY, pp. 161-168.

Khanzode, V.V., Maiti, J., and Ray, P.K. (2011). A methodology for evaluation and monitoring of recurring hazards in coal mining. *Safety Science*, 49, 1172-1179.

Khoo, M.B. and Xie, M. (2009). A study of time-between-events control charts for the monitoring of regularly maintained systems. *Quality and Reliability Engineering International*, 25, 805-819.

Lander, L., Eisen, E.A., Stentz, T.L., Spanjer, K.L., Wendland, B.E., and Perry, M.J. (2011). Near miss reporting system as an occupational injury preventive intervention in manufacturing. *American Journal of Industrial Medicine*, 54, 40-48.

Langley, J. (2004). Challenges for surveillance for injury prevention. *International Journal of Injury Control and Safety Prevention* 11(1), 3-8.

Lucas, J.M. (1985). Counted data CUSUM's. *Technometrics*, 27(2), 129-144.

Montgomery, D. C. (2008). *Introduction to Statistical Quality Control*. Sixth Edition. Hoboken, N.J: John Wiley.

National Institute of Occupational Safety and Health (NIOSH) (2011). "What is Surveillance?" Retrieved from www.cdc.gov/niosh/topics/surveillance/

National Occupational Research Agenda (NORA) (2012). "Surveillance." Retrieved from <http://www.cdc.gov/niosh/programs/surv/goals.html>

Occupational Safety and Health Administration (OSHA) (2013). "Commonly Used Statistics." Retrieved from <http://www.osha.gov/oshstats/commonstats.html>

- Osanaiy, P.A., and Talabai, C.O. (1989). On some non-manufacturing applications of counted data cumulative sum (CUSUM) control chart schemes. *Journal of the Royal Statistical Society, Series D (The Statistician)*, 38(4), 251-257.
- Page, E.S. (1954). Continuous inspection schemes. *Biometrika*, 41, 100-115.
- Peden, M. and Toroyan, T. (2005). Counting road traffic deaths and injuries: poor data should not detract from doing something! *Annals of Emergency Medicine*, 46(2), 158-160.
- Reynolds, M.R., Jr., and Stoumbos, Z.G. (1999). A CUSUM chart for monitoring a proportion when inspecting continuously. *Journal of Quality Technology*, 31(1), 87-108.
- Reynolds, M.R., Jr., and Stoumbos, Z.G. (2000). A general approach to modeling CUSUM charts for a proportion. *IIE Transactions*, 32, 515-535.
- Reynolds, M.R., Jr., and Stoumbos, Z.G. (2004). Control charts and the efficient allocation of sampling resources. *Technometrics*, 46(2), 200-214.
- Schuh, A., Woodall, W.H., and Camelio, J.A. (2013). The effect of aggregation when monitoring a Poisson process. *Accepted for publication in Journal of Quality Technology*.
- Sego, L.H., Woodall, W.H., and Reynolds, M.R., Jr. (2008). A comparison of surveillance methods for small incidence rates. *Statistics in Medicine*, 27, 1225-1247.
- Sethi, D. and Zwi, A. (1999). Traffic crashes another disaster? *European Journal of Public Health*, 9(1), 65-66.
- Sparks, R.S. (2000). CUSUM charts for signalling varying location shifts. *Journal of Quality Technology*, 32, 157-171.
- Sparks, R. (2010). Enhancing road safety through early detection of outbreaks in the frequency of motor vehicle crashes. *Safety Science*, 41, 135-144.
- Szarka, J.L., III and Woodall, W.H. (2011). A review and perspective on surveillance of Bernoulli processes. *Quality and Reliability Engineering International*, 27, 735-752.
- Tsung, F. and Wang, K. (2010). Adaptive charting techniques: literature review and extensions. *Frontiers in Statistical Quality Control*, 9, edited by H.J. Lenz, P.-Th. Wilrich, and W. Schmid, Berlin-Heidelberg, Germany: Springer-Verlag, 19-35.

White, C.H., Keats, J.B., and Stanley, J. (1997). Poisson CUSUM versus c chart for defect data. *Quality Engineering*, 9(4), 673-679.

Woodall, W.H., Grigg, O.A., and Burkom, H.S. (2010). Research issues and ideas on health related surveillance. *Frontiers in Statistical Quality Control*, 9, edited by H.J. Lenz, P.-Th. Wilrich, and W. Schmid, Berlin-Heidelberg, Germany: Springer-Verlag, 145-155.

Woodall, W.H. (2006). The use of control charts in health-care and public-health surveillance. *Journal of Quality Technology*, 38(2), 89-134.

Wu, Z., Liu, Y., Zhen, H., & Khoo, M. (2010). A cumulative sum scheme for monitoring frequency and size of an event. *Quality and Reliability Engineering International*, 26, 541-554.

The following manuscript by Anna Schuh, William H. Woodall, and Jaime A. Camelio has been accepted to appear in the *Journal of Quality Technology*, published by the American Society for Quality. Reprinted with permission from *The Journal of Quality Technology*©2013 American Society for Quality. No further distribution allowed without permission.

4. The Effect of Data Aggregation When Monitoring a Poisson Process

4.1 Abstract

The aggregation of event counts is a common, and often necessary, practice in many applications. When working with large numbers of events, it may be more practical to consider the number of events occurring during time intervals of a given length, rather than the individual interarrival times between successive events. However, when data collection or summarization is done only at the end of aggregation periods, e.g., days, weeks, months, or quarters, there is the potential for significant information loss. In public health and safety applications, this data lag could lead to unnecessary injuries and deaths because of a delay in detecting an increased risk. In our paper, we assume an underlying Poisson process and study the effect of aggregation by investigating the relative performance of the cumulative sum (CUSUM) control chart for monitoring with Poisson-distributed aggregated data and the CUSUM chart based on exponentially-distributed time-between-events data. By comparing steady-state chart performance for varying aggregation periods, we show that the adverse effect of aggregation can be significant.

4.2 Introduction

When the rate of some event of interest is to be monitored, it is often common to aggregate and report numbers of counts over time periods of a specified length. Under the assumption of a Poisson process, these counts are Poisson-distributed and have frequently been monitored with a Shewhart c -chart. These c -charts, however, are known to be less sensitive to small changes in the process mean than cumulative sum (CUSUM) and exponentially-weighted moving average (EWMA) control charts. White et al. (1997) showed that Poisson CUSUM charts more quickly detect changes than Shewhart c -charts and Lucas (1985) suggested the use of Poisson CUSUM charts for monitoring a monthly rate of industrial accidents. Borrer et al. (1998) recommended an EWMA chart for Poisson data.

Lucas (1985) alternatively suggested the use of an exponential CUSUM chart to monitor the times between successive accidents, since the interarrival times of events in a homogeneous Poisson process are exponentially distributed. In addition, Wu et al. (2010) used an exponential CUSUM chart to monitor the times between fatalities resulting from catastrophic accidents. While aggregating counts by week, month, or quarter may provide some simplicity, recording the times between events seems to have an obvious advantage in that there will be no lag in the reporting of information; that is, practitioners will not need to wait until the end of a week or a month, for example, to analyze the process behavior. Benneyan (1998a, 1998b) and Montgomery (2008) also recommended time-between-events monitoring for better efficiency.

Although an important issue, the relative performance of Poisson and exponential CUSUM control charts has only been briefly investigated. Gan (1994) compared the performance of CUSUM charts for exponentially distributed time-between-events data and Poisson-distributed aggregated counts. His results for two aggregation periods showed that the exponential CUSUM chart was more sensitive to changes in the rate of event occurrence than the Poisson CUSUM chart. In addition, Schuh et al. (2013) considered a case study with historical accident data and found that exponential CUSUM charts signaled increases in accident frequency sooner than counts aggregated daily, every three days, weekly, semi-monthly, monthly and quarterly.

The comparison between monitoring methods for aggregated data and non-aggregated data is important because data aggregation has become a common practice in many applications. In the fields of public health and safety, data are often aggregated and reported as numbers of events over time, such as the number of surgical errors in a month or the number of occupational injuries in a week. For instance, the Occupational Safety and Health Administration (OSHA, 2013) collects occupational injury data on an annual basis and the Bureau of Labor Statistics (BLS, 2013) issues a report with this type of data every August or September.

Problems with common public health surveillance methods have been noted and data mining solutions have been proposed for dealing with aggregation. Dubrawski (2011) suggested the use of T-Cube tree-like data representations to store aggregated data and retrieve information in a more timely fashion. In a report issued by the Centers for Disease Control and Prevention (CDC), Burkom et al. (2004) investigated the role of data aggregation in biosurveillance data from a Defense Advanced Research Project Agency (DARPA) project. The authors

recommended Bayesian belief networks for the temporal and spatial monitoring of epidemiological outbreaks. While these researchers discussed issues and types of aggregation and addressed aggregation for data storage, we are more interested in studying the effect of data aggregation over time, such as the annual, monthly, or weekly reporting of events.

The effect of data aggregation in control charting applications has been investigated for some common situations in manufacturing when monitoring normally distributed data. Reynolds and Stoumbos (2004a) showed that with a constant sampling rate smaller samples, i.e. with less aggregation, were much better at detecting large shifts in the mean of the process. Similarly, Reynolds and Stoumbos (2004b) compared the use of concentrated sampling and dispersed sampling, reporting that non-aggregated samples of individual measurements resulted in the best performance.

The CUSUM charts with data from Bernoulli and geometric distributions, the counterparts of the Poisson and exponential distributions for binary data, have also been studied. The Bernoulli CUSUM chart considers the individual inspection of items and the result (conforming or non-conforming) after each inspection. Geometric charts are based on the number of conforming items between successive nonconforming items. Reynolds and Stoumbos (1999) found that the Bernoulli CUSUM chart more quickly detected an increase in the proportion of nonconforming items than standard Shewhart charts based on the binomial distribution resulting from aggregation. Building on this work, Reynolds and Stoumbos (2000) compared the performance of Bernoulli CUSUM charts to that of binomial CUSUM charts for aggregated samples, finding that Bernoulli CUSUM charts showed better overall performance, especially for detecting large increases in the rates of nonconforming items. Szarka and Woodall (2011) further discussed this topic in their review of Bernoulli-based charts.

Our paper extends these previous investigations of the effect of data aggregation by more thoroughly exploring the relative performance of exponential CUSUM and Poisson CUSUM control charts with some longer aggregation periods. We expand upon the work of Gan (1994) by considering the steady-state performance of the charts, simulating a randomly occurring sustained shift in the rate at some point after monitoring begins. We also consider more commonly used aggregation periods in public health and safety, i.e., weekly, bi-weekly, and monthly, for one special case.

4.3 Poisson and Exponential CUSUM Control Charts

In this section we introduce the background and structure of the Poisson and exponential CUSUM charts. We assume a homogeneous Poisson process with an average time between events of μ time units. Thus the interarrival times are independently distributed exponential random variables with mean μ . For an aggregation period of length d time units, counts are independent and Poisson distributed with a mean of $\mu = d/\mu$. The time unit considered can vary depending on the application. We assume that the in-control value of the parameter is $\mu_0=1$ and that we wish to detect only decreases in the average time between events. Our results, however, are generalizable. For example, if events occur at an average rate of 4 per 28 day period, then this is equivalent to one event per week. If we consider an aggregation period of seven time units, then this would correspond to aggregating the data over a 49 day period. Thus, it is only necessary to consider the $\mu_0=1$ case.

4.3.1 Poisson CUSUM Control Charts

In a traditional Shewhart c -chart, counts of events are monitored for unusual behavior. A Phase I baseline of data is used to establish an upper control limit and the counts are then monitored in real time during Phase II. When the number of events exceeds the established control limit, the chart signals and the process is considered to be out-of-control. In manufacturing settings, the process may be stopped, the source of the increased mean is investigated and fixed, and the monitoring process begins again. In public health and safety applications, the process is not so easily stopped. Therefore, a safety monitoring process should be restarted following the implementation of a new hazard control, as Schuh et al. (2013) suggested.

Page (1954) introduced the CUSUM chart, which instead incorporates historical data to more effectively detect parameter step changes. Since an aggregated number of events in defined time periods is Poisson-distributed, a Poisson CUSUM chart is appropriate for monitoring such counts. At time i , the Poisson CUSUM statistic used to detect an increase in the average number of events is given by

$$S_{\text{POISSON}, i} = \max (0, S_{\text{POISSON}, i-1} + X_i - k_{\text{POISSON}}), \quad (1)$$

where X_i is the observed count at time i and k_{POISSON} is the reference parameter. This reference parameter is given by Lucas (1985) to be

$$k_{\text{POISSON}} = \frac{\tilde{\mu}_1 - \tilde{\mu}_0}{\ln(\tilde{\mu}_1) - \ln(\tilde{\mu}_0)} \quad (2)$$

where μ_0 is the in-control mean number of events and μ_1 is the mean number of events to be detected quickly. The value $S_{\text{POISSON}, 0}$ is usually set to zero. The value μ_0 is typically estimated as the Phase I mean, but we assume the value is known. The chart signals when $S_{\text{POISSON}, i} > h$. For all of the examples to be discussed in this paper, h is set to give a desired in-control average time to signal (ATS) value.

4.3.2 Exponential CUSUM Control Charts

An alternative to monitoring aggregated numbers of events over time is to monitor the times between successive adverse events. Thus, information will be incorporated as it is obtained, instead of only updating the monitoring procedure at the end of specified time periods. An increase in the average time between events is desirable and a decrease should be detected and investigated as quickly as possible. The exponential CUSUM statistic at time i is given by

$$C_{\text{EXP}, i} = \max(0, C_{\text{EXP}, i-1} + k_{\text{EXP}} - Y_i) \quad (3)$$

where Y_i is the time between the current event i and the previous event $i-1$. The value $C_{\text{EXP}, 0}$ is usually set to 0. If λ_1 is the rate of adverse events to be detected quickly and λ_0 is the in-control rate, the reference parameter is given by

$$k_{\text{EXP}} = \frac{\lambda_1 \lambda_0}{\lambda_1 - \lambda_0} \ln\left(\frac{\lambda_1}{\lambda_0}\right). \quad (4)$$

Lucas (1985) converted these rates to counts per time period, implicitly assuming $d=1$, and defined the exponential CUSUM reference parameter as

$$k_{\text{EXP}} = \frac{\ln(\tilde{\mu}_1) - \ln(\tilde{\mu}_0)}{\tilde{\mu}_1 - \tilde{\mu}_0}. \quad (5)$$

Equations (4) and (5) produce the same value for the reference parameter.

4.4 Relative Performance of Poisson and Exponential CUSUM Control Charts

In this section we will investigate the relative performance of Poisson and exponential CUSUM charts for various levels of aggregation using computer simulation.

4.4.1 Zero-state Performance

As mentioned in Section 4.2, Gan (1994) briefly studied the relative performance of Poisson and exponential CUSUM control charts. Gan's zero-state ATS values are given in Table 4.1. The CUSUM charts he considered were optimal for detecting average interarrival times changing from $\lambda_0=1$ to $\lambda=0.5$. Therefore, the charts will be optimal in detecting an increase from an average of one event per time period to two ($1/0.5 = 2$). The Poisson charts, represented by POIS1 and POIS10, use aggregation time periods of length 1 and 10 time units, respectively. When considering a public health example, it may be helpful to consider the time period as one day – so the in-control ATS values for $\lambda=1$ indicate that a signal will occur, on average, after about 204 or 1112 days when the process is in-control, depending on the desired in-control ATS. The reference values (k) and control limits (h) in Table 4.1 for each chart were given by Gan (1994).

We agree with Gan's conclusion that the zero-state exponential CUSUM chart is more sensitive to decreases in average interarrival times (increases in the rate of event occurrence) than the Poisson CUSUM charts based on aggregation time intervals of 1 and 10. This is most clearly the case for large decreases in λ . One should note, however, that these could correspond to relatively low levels of aggregation

		POIS10	POIS1	EXP	POIS10	POIS1	EXP
	<i>k</i>	14.429	1.44	0.693	14.429	1.44	0.693
	<i>h</i>	1.429	4.68	3.297	3.714	7.08	4.927
1		205.0	204.3	204.3	1113.8	1112.2	1112.2
0.95		143.7	144.9	144.8	653.6	654.6	653.5
0.90		101.2	102.6	102.4	383.2	384.0	382.9
0.85		71.8	72.8	72.3	226.2	226.3	225.3
0.80		51.4	51.8	51.2	135.6	135.3	134.4
0.75		37.4	37.2	36.5	83.5	83.0	82.1
0.70		27.8	27.0	26.2	53.5	52.7	51.8
0.65		21.2	19.9	19.1	36.0	35.0	34.0
0.60		16.7	14.9	14.0	25.5	24.2	23.3
0.55		13.7	11.3	10.5	19.1	17.4	16.5
0.50		11.9	8.7	7.9	15.0	12.9	12.1
0.45		10.8	6.8	6.0	12.5	9.8	9.0
0.40		10.2	5.4	4.6	11.0	7.5	6.8
0.35		10.0	4.3	3.5	10.2	5.8	5.2
0.30		10.0	3.4	2.7	10.0	4.5	3.9
0.25		10.0	2.6	2.0	10.0	3.5	2.9
0.20		10.0	2.1	1.4	10.0	2.6	2.1

Table 4.1: Gan (1994) zero-state ATS for optimal (in detecting shift to $=0.5$) Poisson CUSUM and exponential CUSUM control charts

4.4.2 Steady-State Performance

In the zero-state performance comparisons, like that presented by Gan (1994), it is assumed that the process is already in an out-of-control state when monitoring begins or that the control chart statistic is at its initial value when a sustained shift in the parameter occurs. In Gan's zero-state results, the ATS values were calculated as the average time between the beginning of the monitoring period and an out-of-control signal. For the Poisson aggregated charts, the average run length (ARL) was calculated as the number of time periods until a signal. The ARL can then be converted to an ATS value by multiplying the number of time periods by the length of each time period. For example, an ARL of 10 ten-day time periods would be the same as an ATS of 100 days to signal. Only ATS values are reported in Table 4.1, to allow for easier comparisons among the Poisson CUSUM charts and the exponential CUSUM chart.

In applications, it is not reasonable to assume that a process is already out-of-control when monitoring begins or that the control chart statistic is at its initial value when any shift occurs.

Instead, it is expected that a period of in-control behavior will be followed by a randomly occurring shift to out-of-control behavior. It is of interest to know how soon a signal occurs with each type of chart following this random shift, as this will indicate the chart's value in practical use. This metric, the average time between the occurrence of the shift and a signal, is known as a chart's steady-state ATS performance and sometimes referred to as the conditional expected delay. Table 4.2 shows the steady-state ATS values for the first set of charts considered in Table 4.1, based on 100,000 simulated charts. Again it is clear that the exponential CUSUM chart exhibits better performance than either of the Poisson charts considered.

		POIS10	POIS1	EXP
	<i>k</i>	14.429	1.44	0.693
	<i>h</i>	1.429	4.68	3.297
1		204.5	203.8	204.9
0.95		144.8	143.8	142.8
0.9		102.4	102.4	101.2
0.85		73.2	70.6	70.6
0.8		53.2	50.2	50.2
0.75		39.1	35.9	35.9
0.7		29.4	26.0	25.6
0.65		22.6	18.9	18.5
0.6		17.8	14.0	13.6
0.55		14.5	10.6	10.1
0.5		12.1	8.2	7.6
0.45		10.4	6.3	5.9
0.4		9.2	5.0	4.5
0.35		8.3	3.9	3.5
0.3		7.6	3.1	2.6
0.25		7.0	2.4	2.0
0.2		6.5	1.9	1.4

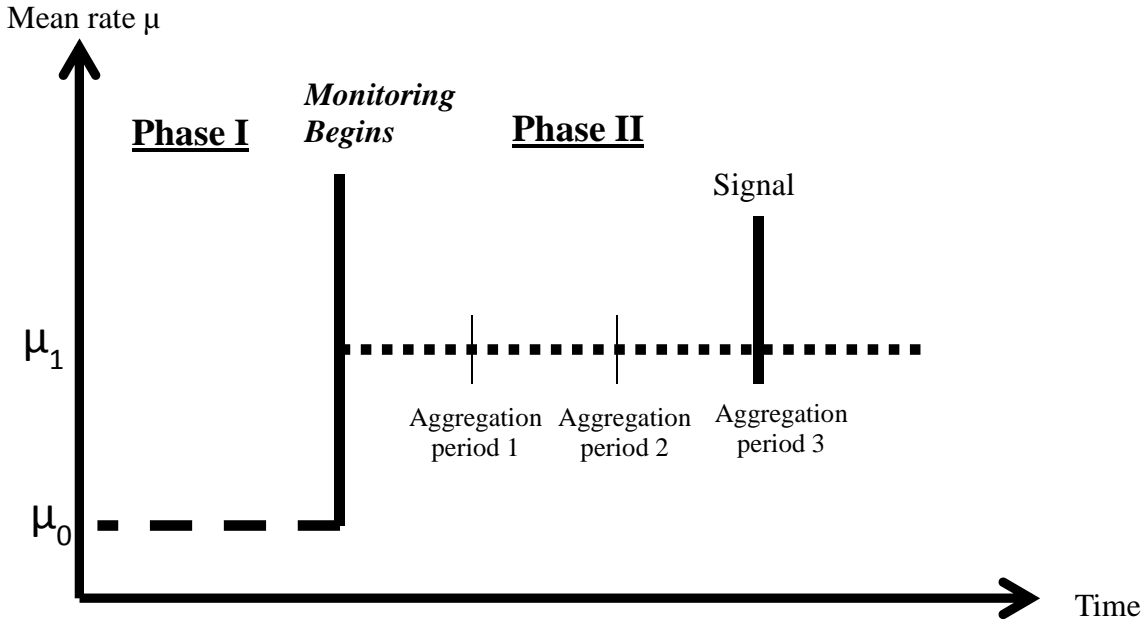
Table 4.2: Steady-state average time to signal for optimal (in detecting a sustained shift to $\lambda = 0.5$) Poisson and exponential CUSUM control charts with respect to an increase in the rate of occurrences of events

The Poisson chart ATS values in Table 4.2 were calculated by designating a specified number of time periods (50) as an in-control period, referred to as the warm-up period. All simulated events in the warm-up period follow a Poisson distribution with parameter 1 or 10, as the case may be. Since the randomly-occurring shift to out-of-control behavior will not necessarily occur at the

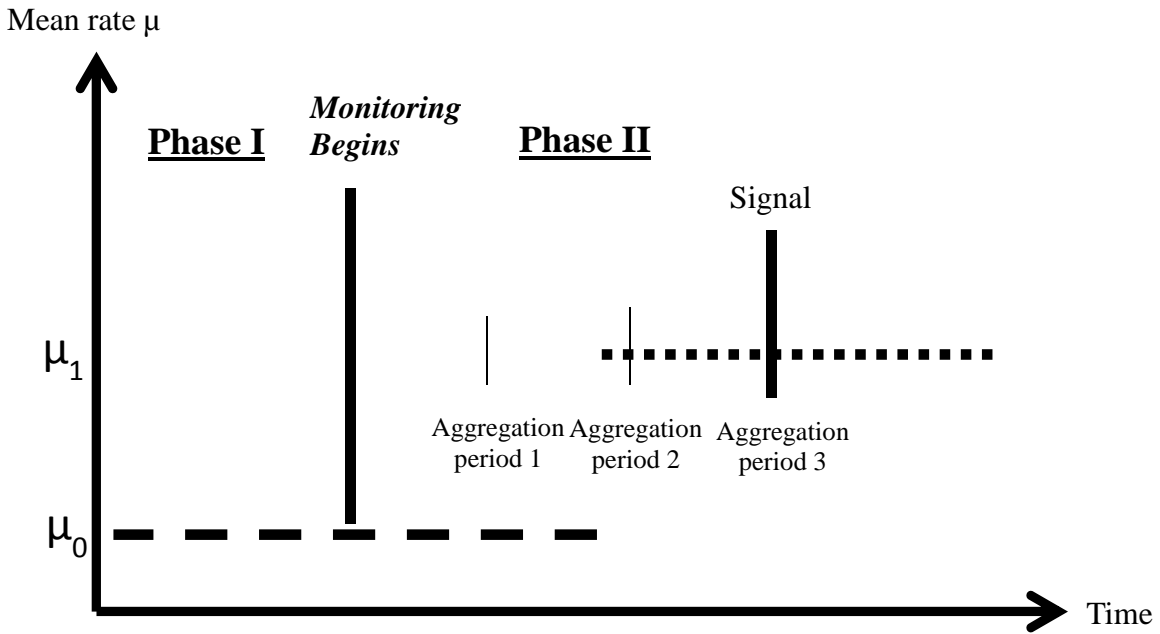
beginning of an aggregated time period (i.e., the beginning of a month or a week), the shift was assumed to occur at a uniformly distributed random time during the first time period following the warm-up period. Therefore, the number of events simulated for this time period is partially affected by the in-control parameter value μ_0 and partially affected by the out-of-control value μ_1 . We assume that the out-of-control shift will be sustained, so each time period thereafter will have a simulated number of events with Poisson parameter μ_1 .

Likewise, for the exponential time-between-events chart, the shift will likely occur between events, not immediately following an event. Therefore, the time between the last event when the process was in-control and the first event when the process was out-of-control will be affected by λ_0 for a random percentage of the time and λ_1 for the rest. Again, since the shift is sustained, the time between the succeeding events will be distributed with the parameter λ_1 . We assumed that the time of the shift was uniformly distributed in the interval $(10^{-1}\lambda_0, 20^{-1}\lambda_0)$.

The distinction between the simulations for zero-state and steady-state conditions is illustrated in Figures 4.1 and 4.2. In Figure 4.1(b), one can see that the count for one aggregation period is influenced by the values of both μ_0 and μ_1 . For simplicity of presentation, it is assumed here that the shift in the parameter occurs in the second aggregation period. Similarly, in Figure 4.2(b), one time between events is influenced by both λ_0 and λ_1 . Szarka and Woodall (2012) described a similar framework when monitoring with Bernoulli data.

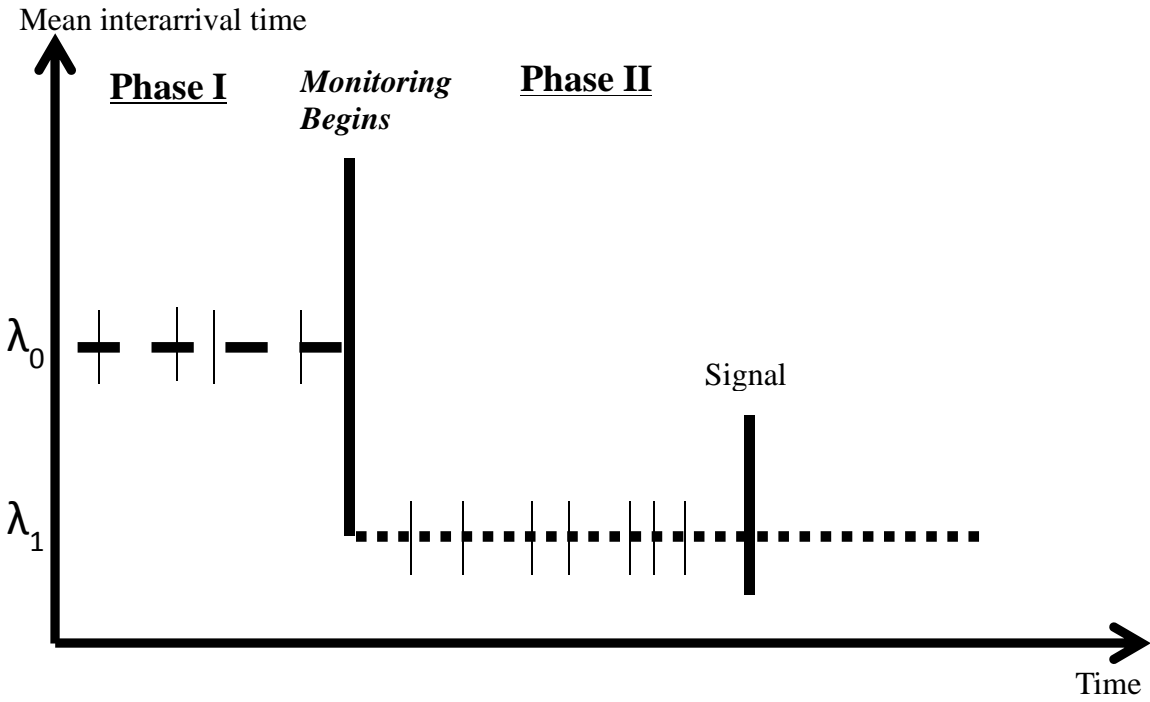


(a)

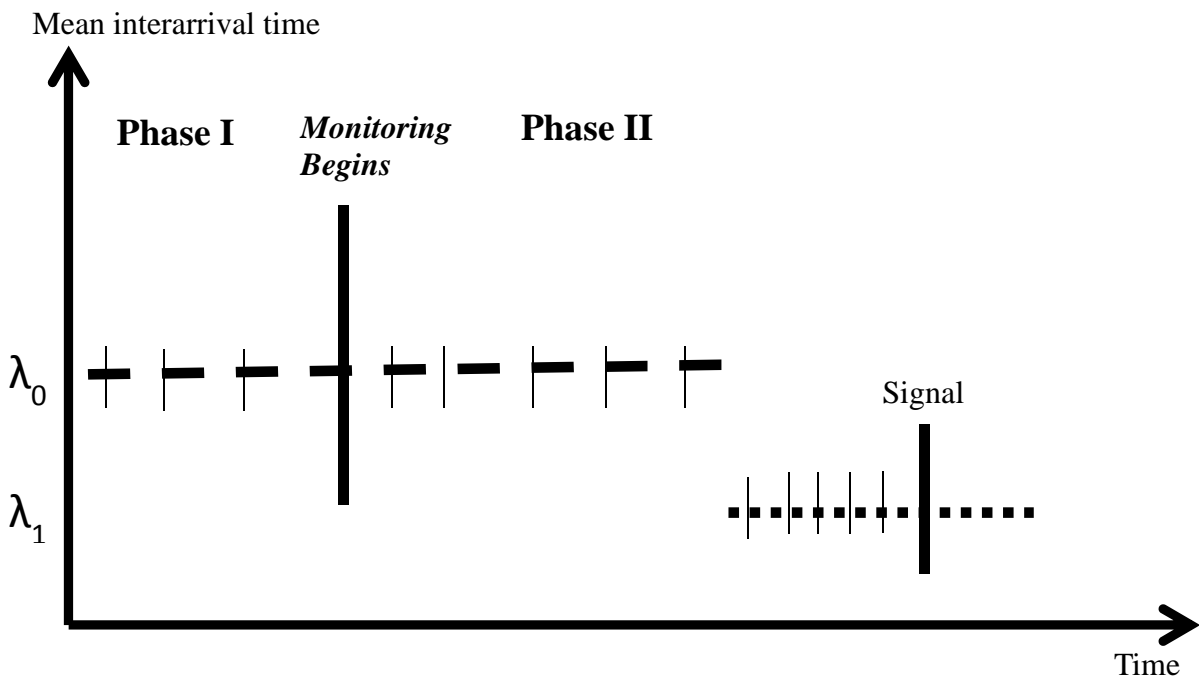


(b)

Figure 4.1: Zero-state (a) and steady-state (b) scenarios for monitoring with aggregated Poisson counts



(a)



(b)

Figure 4.2: Zero-state (a) and steady-state (b) scenarios monitoring with exponential interarrival times

Because the steady-state Poisson simulations allow a partially out-of-control period for the first time period after the warm-up, it is possible that the shift will be detected after only this partial period. This is most likely to happen when detecting large decreases in λ and is why some of the steady-state ATS values shown in Table 4.2 are less than 10 for the POIS10 chart.

4.4.3 Extension to Other Aggregation Periods

When considering the use of exponential or Poisson CUSUM charts in public health applications, it may be helpful to consider longer aggregation periods. For instance, it may be more common to record an aggregated number of surgical failures in a week or the number of industrial accidents in a month, which may not correspond to the one or ten time periods Gan (1994) considered. Therefore, in Tables 4.3 and 4.4, we compare the ATS performance of exponential CUSUM charts to Poisson CUSUM charts based on aggregation time intervals of 1, 7, 14, and 30 time periods in order to simulate daily, weekly, bi-weekly, and monthly aggregation with $\tau_0=1$ day. Table 4.3 shows similar in-control performance for the charts considered and displays their steady state performance. The reference values were calculated using Equations (2) and (4), so these values for the POIS1 and EXP charts vary slightly from Gan's (1994) reference values given in Table 4.1.

The CUSUM statistics are discrete for Poisson-distributed data. The zero-state in-control ATS values follow a step function with respect to the control limit, as described by Han et al. (2009). Poisson charts with different reference parameters, like those given in Table 4.3, have different values along these step functions. The zero-state parameters suggested by Gan (1994) for charts with data aggregated over 1 and 10 time periods produce ATS values that do not correspond to values obtained along the step functions for aggregated data over 7, 14 and 30 time periods. Therefore, the in-control ATS values used in Table 4.3 were carefully chosen so that they were all approximately the same. Since more charts are being compared than in Gan (1994), this was not a trivial undertaking; a compromise was necessary to allow somewhat higher in-control ATS values for the POIS14 and POIS30 charts. Furthermore, it should be noted that the in-control ATS values must be larger when comparing charts with longer aggregation periods, in order to find a control limit greater than zero for the chart corresponding to the longest aggregation period. Even with a relatively high desired in-control ATS of around 4400, the POIS30 chart requires a control limit close to zero to approximately match it.

		POIS30	POIS14	POIS7	POIS1	EXP
	<i>k</i>	43.2809	20.1977	10.0989	1.4427	0.6931
	<i>h</i>	0.7191	4.8023	6.901	9.0182	6.2877
1		4,790.1	4,482.2	4,418.3	4,420.0	4,416.8
0.95		2,127.4	2,227.9	2,296.7	2,182.2	2,178.7
0.9		976.3	1,090.8	1,157.5	1,062.6	1,069.3
0.85		468.4	542.7	579.6	530.4	528.0
0.8		239.1	274.6	294.0	265.6	268.6
0.75		133.6	148.5	154.8	140.3	141.6
0.7		82.5	85.5	86.7	80.9	79.1
0.65		52.9	53.6	53.0	48.1	47.8
0.6		42.3	36.7	35.0	31.3	31.0
0.55		34.3	27.1	24.8	21.6	21.3
0.5		29.7	21.2	18.6	15.6	15.3
0.45		26.6	17.4	14.4	11.7	11.3
0.4		24.4	14.9	11.6	9.0	8.5
0.35		22.6	13.1	9.6	6.9	6.5
0.3		21.0	11.8	8.1	5.3	4.9
0.25		19.6	10.7	6.9	4.0	3.7
0.2		18.5	9.7	6.0	3.0	2.6

Table 4.3: Steady-state performance for optimal (in detecting shift to $\lambda = 0.5$) Poisson and exponential CUSUM control charts with respect to an increase in the rate of occurrences of events

It is important to recognize again that we are assuming each of these aggregation periods represent units of days, in order to better conceptualize the specific impact for public health and safety applications. For instance, when considering an increase in event frequency from $\lambda_0 = 1$ to $\lambda = 0.2$ in Table 4.3, we can see that a signal would be given on average after 6.0 time units when monitoring with the POIS7 chart, compared to 9.7 time units when using the POIS14 chart. We assume that we are aggregating over seven days and fourteen days, respectively, since these are common aggregation periods in public health – and, therefore, we interpret the values in the table to indicate that a signal would be expected 3.7 days sooner when aggregating weekly instead of bi-weekly for this particular example. However, these aggregation periods could just as easily be considered in terms of any time unit: seconds, hours, weeks, months, etc. If the time period were weeks instead of days, for example, we would be aggregating over a 7, 14, or 30 week period. Likewise, we assume that the effective time unit is that for which the expected count equals one.

If the expected count per time period was two, the POIS30 chart would be based on aggregated counts for 15 time periods.

Since our overall goal is to reduce the number of observed adverse events, Table 4.4 reiterates the impact of data aggregation by showing the number of additional adverse events that would be expected to occur as a result of using any chart instead of the most efficient chart for detecting a given increase in event frequency. For example, when detecting an increase in the frequency of events to $\lambda = 0.2$, using the POIS1 chart instead of the EXP chart would result in the occurrence of an expected 2.0 additional adverse events before a signal. Steiner and Jones (2010) considered a similar metric when comparing control chart performance for the monitoring of surgical deaths.

		POIS30	POIS14	POIS7	POIS1	EXP
	<i>k</i>	43.2809	20.1977	10.0989	1.4427	0.6931
	<i>h</i>	0.7191	4.8023	6.901	9.0182	6.2877
0.95		---	105.8	178.2	57.7	54.0
0.9		---	127.2	201.3	95.9	103.3
0.85		---	87.4	130.8	72.9	70.1
0.8		---	44.4	68.6	33.1	36.9
0.75		---	19.9	28.3	8.9	10.7
0.7		8.0	9.1	10.9	2.6	---
0.65		7.8	8.9	8.0	0.5	---
0.6		18.8	9.5	6.7	0.5	---
0.55		23.6	10.5	6.4	0.5	---
0.5		28.8	11.8	6.6	0.6	---
0.45		34.0	13.6	6.9	0.9	---
0.4		39.8	16.0	7.8	1.3	---
0.35		46.0	18.9	8.9	1.1	---
0.3		53.7	23.0	10.7	1.3	---
0.25		63.6	28.0	12.8	1.2	---
0.2		79.5	35.5	17.0	2.0	---

Table 4.4: Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency

From Tables 4.3 and 4.4, it is clear that for large decreases in average interarrival times, the exponential chart provides the best performance, as we expected. However, for small decreases in λ values, charts with longer aggregation periods begin to perform more efficiently than those with shorter aggregation periods. For detecting shifts to λ values between .7 and .95, the chart

based on aggregating 14 time periods performs better than the one based on aggregating 7 time periods. When detecting decreases in λ to values between .75 and .95, the aggregated Poisson chart based on 30 time periods outperforms the exponential CUSUM chart. Similar results were found by Reynolds and Stoumbos (2004a) for the Bernoulli/binomial case. However, since the charts in Table 4.4 are designed to be optimal for detecting a decrease in interarrival times from 1.00 to 0.5, it is not reasonable to expect that the exponential chart will have optimal performance for detecting smaller decreases. By only comparing two Poisson charts with smaller aggregation periods to the exponential chart in the zero-state, Gan (1994) did not observe the phenomenon of better performance with longer aggregation periods. Even when the steady-state performance was considered for the charts in Table 4.2, this somewhat counterintuitive effect went unnoticed. It should be kept in mind, however, that if the small shifts are considered important, the charts should be tuned appropriately to detect them quickly.

The expected numbers of additional adverse events from Table 4.4 are also plotted in Figure 4.3, in order to more easily compare the charts. The graph further emphasizes the observation that the EXP chart is best for detecting larger decreases in λ .

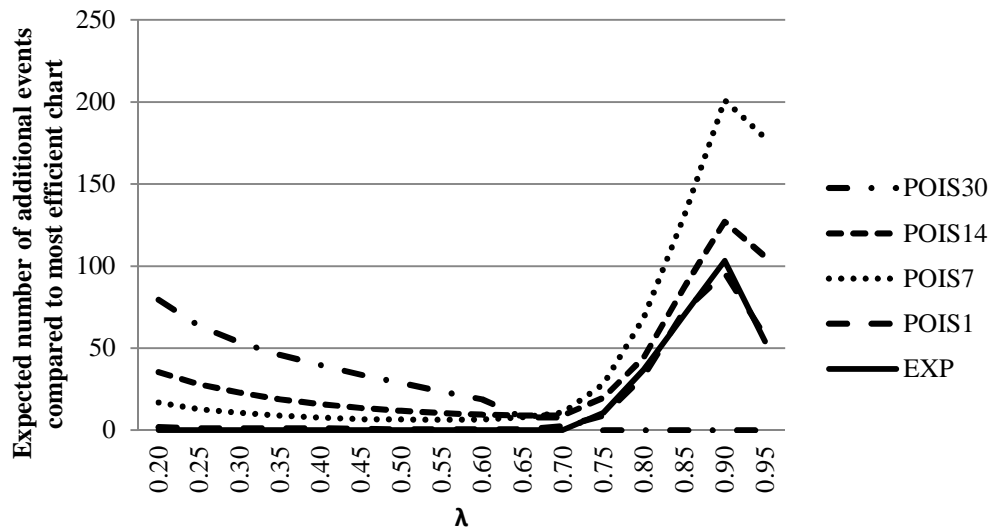


Figure 4.3: Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency, based on data from Table 4.4

4.4.4 Longer Aggregation Periods

One will notice that the values of the control limits for the Poisson charts in Table 4.3 are tending toward zero as the amount of aggregation increases. In fact, as the length of the aggregation period increases, the control limits decrease and the CUSUM charts become more like Shewhart charts. The charts are designed to detect a decrease in λ from 1.00 to 0.50, which is the equivalent of detecting a doubling of the process mean from one event per time period to two. Since the process mean is the same as the variance for a Poisson process, the POIS1 chart will be detecting a shift in the average count value from one to two, which is the same as a one standard deviation increase. However, when detecting a shift from an average of 30 counts to 60 counts with the POIS30 chart, this is a difference of 5.5 standard deviations. As the shift size increases, the CUSUM chart's benefit over a Shewhart chart decreases. A practitioner may prefer to use a Shewhart chart because they are easier to design.

Since it is common for some agencies and industries to aggregate data over long time periods, it is important to consider the performance of control charts designed to monitor data collected over these longer aggregation periods. In order to study the performance of Poisson CUSUM charts based on longer aggregation periods without tending toward Shewhart behavior so quickly, one could design the charts to detect smaller decreases in λ . For instance, if the charts were designed to be optimal for detecting a shift in λ from 1.00 to 0.75, they would be optimal for detecting a one-third increase in the process mean count instead of a doubling. When the charts are designed to detect smaller decreases in interarrival times, longer aggregation periods can be studied because the CUSUM control limits will tend toward zero at a slower rate with increasing levels of aggregation. Table 4.5 compares four Poisson charts to the exponential chart, this time all designed to be optimal in detecting a shift to $\lambda = 0.75$. Longer aggregation periods are considered this time; the performance of Poisson charts with annual, semi-annual, quarterly, and monthly Poisson aggregation periods, given an in-control rate of one per day, are compared to the exponential chart. Again, it was necessary for the in-control ATS values to be extremely high in order to have a control limit greater than zero for the POIS365 chart. As a result of this high in-control ATS, the number of simulated charts has been reduced to 10,000 to estimate the ATS values in this table.

		POIS365	POIS180	POIS90	POIS30	EXP
	<i>k</i>	422.9206	208.5636	104.2818	34.7606	0.8630
	<i>h</i>	1.1895	16.6746	21.8022	26.4982	27.5151
1		<i>311,360.0</i>	<i>313,840.0</i>	<i>311,230.0</i>	<i>316,390.0</i>	<i>312,150.0</i>
0.95		<i>17,346.0</i>	<i>26,688.0</i>	<i>28,086.0</i>	<i>26,896.0</i>	<i>26,393.0</i>
0.9		<i>2,218.5</i>	<i>3,243.4</i>	<i>3,232.0</i>	<i>3,104.0</i>	<i>3,059.6</i>
0.85		<i>708.4</i>	<i>772.8</i>	<i>731.8</i>	<i>688.8</i>	<i>686.4</i>
0.8		<i>442.9</i>	<i>363.4</i>	<i>322.0</i>	<i>296.7</i>	<i>288.2</i>
0.75		<i>362.2</i>	<i>248.2</i>	<i>203.3</i>	<i>174.2</i>	<i>167.5</i>
0.7		<i>320.6</i>	<i>199.6</i>	<i>148.7</i>	<i>119.8</i>	<i>111.6</i>
0.65		<i>291.5</i>	<i>174.3</i>	<i>118.1</i>	<i>90.0</i>	<i>80.9</i>
0.6		<i>271.4</i>	<i>156.6</i>	<i>100.5</i>	<i>71.7</i>	<i>61.4</i>
0.55		<i>253.7</i>	<i>144.9</i>	<i>89.6</i>	<i>58.3</i>	<i>47.4</i>
0.5		<i>240.5</i>	<i>134.7</i>	<i>80.9</i>	<i>48.9</i>	<i>37.2</i>
0.45		<i>229.6</i>	<i>126.7</i>	<i>74.1</i>	<i>41.4</i>	<i>29.6</i>
0.4		<i>221.8</i>	<i>120.3</i>	<i>68.5</i>	<i>35.7</i>	<i>23.4</i>
0.35		<i>212.4</i>	<i>113.9</i>	<i>64.5</i>	<i>31.7</i>	<i>18.6</i>
0.3		<i>208.0</i>	<i>109.3</i>	<i>59.6</i>	<i>28.2</i>	<i>14.5</i>
0.25		<i>202.6</i>	<i>104.9</i>	<i>56.4</i>	<i>25.1</i>	<i>11.1</i>
0.2		<i>197.7</i>	<i>101.6</i>	<i>53.7</i>	<i>22.6</i>	<i>8.2</i>

Table 4.5: Steady-state performance for optimal (in detecting shift to $\lambda=0.75$) Poisson and exponential CUSUM control charts with respect to an increase in the rate of occurrences of events

As we noted in Table 4.3, the Poisson charts in Table 4.5 with the longer aggregation periods have better ATS performance for detecting small decreases in interarrival times. However, the exponential chart again shows much quicker detection of larger decreases, including the targeted decrease in average interarrival times to $\lambda=0.75$. To further demonstrate this, Table 4.6 shows the additional adverse events expected to result from use of less effective charts in Table 4.5.

		POIS365	POIS180	POIS90	POIS30	EXP
	<i>k</i>	422.9206	208.5636	104.2818	34.7606	0.8630
	<i>h</i>	1.1895	16.6746	21.8022	26.4982	27.5151
0.95		---	9,833.7	11,305.3	10,052.6	9,523.2
0.9		---	1,138.8	1,126.1	983.9	934.6
0.85		25.9	101.6	53.4	2.8	---
0.8		193.4	94.0	42.3	10.6	---
0.75		259.6	107.6	47.7	8.9	---
0.7		298.6	125.7	53.0	11.7	---
0.65		324.0	143.7	57.2	14.0	---
0.6		350.0	158.7	65.2	17.2	---
0.55		375.1	177.3	76.7	19.8	---
0.5		406.6	195.0	87.4	23.4	---
0.45		444.4	215.8	98.9	26.2	---
0.4		496.0	242.3	112.8	30.8	---
0.35		553.7	272.3	131.1	37.4	---
0.3		645.0	316.0	150.3	45.7	---
0.25		766.0	375.2	181.2	56.0	---
0.2		947.5	467.0	227.5	72.0	---

Table 4.6: Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency

Figure 4.4 graphs the expected additional number of adverse event values from Table 4.6, for the detection of λ in the range from 0.2-0.85, in order to better visualize the benefit of using charts with smaller aggregation periods for the detection of larger increases in interarrival times. Since the EXP chart is the most effective of our charts for λ in the range 0.2-0.85, it appears as a horizontal line at 0 in Figure 4.4, and the expected number of additional adverse events resulting from the use of Poisson charts increases with increasing levels of aggregation. Since these charts were not designed to detect very small increases, the results for λ in the range 0.9-0.95 have been omitted from Figure 4.4.

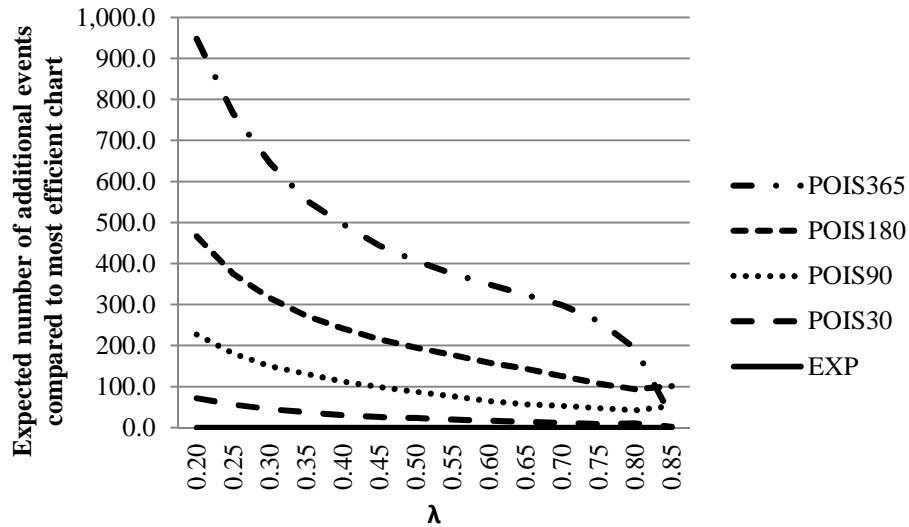


Figure 4.4: Number of additional adverse events expected for each less-effective chart when detecting various increases in event frequency, based on decreases to $\lambda=0.2-0.85$ from Table 4.6. As we have discussed, for high levels of aggregation where the control limit is nearing zero, the Poisson CUSUM chart is expected to show similar performance to a Shewhart c -chart. Table 4.7 compares the ATS performance of Shewhart charts to the Poisson CUSUM charts with the longest aggregation periods from Tables 3 and 5, POIS30 and POIS365. We notice that Shewhart charts offer comparable performance to Poisson CUSUM charts for these longer aggregation periods, but the exponential chart is still preferred for the detection of large decreases in λ .

When considering data aggregated over very long time periods, a practitioner may consider implementing a Shewhart chart instead of a Poisson CUSUM chart. Ideally, however, current processes using these extremely long aggregation periods should be altered such that collected data are aggregated more frequently, or time-between-events data are monitored instead. This paper has shown that exponential CUSUM charts for time-between-events data have the best overall detection performance, and Poisson CUSUM charts using short aggregation periods are preferred if temporal data aggregation is necessary.

		POIS30	Shewhart30	POIS365	Shewhart365
	<i>k</i>	43.2809	N/A	422.9206	N/A
	<i>h</i>	0.7191	44.0	1.1895	424.0
	1	4,790.1	4,726.3	311,360.0	310,740.0
	0.95	2,127.4	2,138.2	17,346.0	17,180.0
	0.9	976.3	982.7	2,218.5	2,257.8
	0.85	468.4	473.2	708.4	710.0
	0.8	239.1	242.5	442.9	439.3
	0.75	133.6	136.0	362.2	359.1
	0.7	82.5	84.2	320.6	320.5
	0.65	52.9	56.6	291.5	292.6
	0.6	42.3	42.5	271.4	271.9
	0.55	34.3	34.5	253.7	254.0
	0.5	29.7	29.7	240.5	240.1
	0.45	26.6	26.6	229.6	231.1
	0.4	24.4	24.4	221.8	221.6
	0.35	22.6	22.6	212.4	213.4
	0.3	21.0	21.0	208.0	207.9
	0.25	19.6	19.7	202.6	202.5
	0.2	18.5	18.6	197.7	198.8

Table 4.7: Steady-state performance for Poisson CUSUM charts (POIS30 from Table 4.4, and POIS365 from Table 4.6) and Shewhart control charts, with respect to an increase in the rate of occurrences of events

4.5 Conclusions

The need for better comparisons between aggregated and non-aggregated data reporting methods for public health and safety monitoring is important. In our paper, previous work in the comparison of Poisson and exponential CUSUM charts was expanded through an analysis of steady-state chart performance, which better represents the actual occurrence of adverse events in public health and safety. The exponential chart was observed to have better overall performance than Poisson charts based on aggregated data. Longer aggregation periods were investigated by designing the charts to detect smaller decreases in interarrival times. The expected number of additional adverse events when less efficient charts are used was also calculated to further motivate the need for real-time monitoring. Finally, more straightforward Shewhart charts were found to have comparable ATS performance to Poisson CUSUM charts when the length of the aggregation period is very long. However, these processes should ideally

be restructured to allow for more frequent data collection, since the exponential chart is preferred.

Although we have considered the monitoring of a homogeneous Poisson process in our paper, our approach can be generalized to study the effect of aggregation on seasonal data, autocorrelated data, and data that includes other effects. It can also be extended to risk-adjusted monitoring in healthcare. Research is needed on the effect of aggregation on the detection of rate decreases as well as on its impact on detecting transient process changes. We strongly encourage practitioners to study the effect of aggregation in their applications. Even though it is sufficient to consider only the $\theta_0=1$ case, it is not practical for us to consider all of the aggregation levels that may be relevant in particular applications.

As a final note, it is important in many applications to investigate each serious adverse event even if there is no evidence of a process change. Such investigations and the implementation of process improvements do not preclude the use of control charts to determine whether or not the rate of adverse events has increased or decreased. The detection of increases in the average time between events would indicate the success of a process improvement initiative.

4.6 References

Benneyan, J.C. (1998a). “Statistical Quality Control Methods in Infection Control in Hospital Epidemiology, Part I: Introduction and Basic Theory.” *Infection Control and Hospital Epidemiology* 19, pp. 194-219.

Benneyan, J.C. (1998b). “Statistical Quality Control Methods in Infection Control in Hospital Epidemiology, Part II: Chart Use, Statistical Properties, and Research Issues.” *Infection Control and Hospital Epidemiology* 19(4), pp. 265-283.

Borror, C.M.; Champ, C.W.; and Rigdon, S.E. (1998). “Poisson EWMA Control Charts.” *Journal of Quality Technology* 30(4), pp. 352-361.

Bureau of Labor Statistics (BLS) (2013). Injuries, Illnesses and Fatalities. www.bls.gov

Burkom, H. S., Elbert, Y., Feldman, A., and Lin, J. (2004). “Role of Data Aggregation in Biosurveillance Detection Strategies with Applications from ESSENCE”. *Morbidity and Mortality Weekly Report* 53 (Suppl), 67-73.

Dubrawski, A. (2011). “The Role of Data Aggregation in Public Health and Food Safety Surveillance”, Chapter 9 in *Biosurveillance: Methods and Case Studies*, edited by T. Kass-Hout and X. Zhang, Boca Raton, FL: Taylor and Francis Group, LLC.

Gan, F.F. (1994). "Design of Optimal Exponential CUSUM Control Charts." *Journal of Quality Technology* 26(2), pp. 109-124.

Han, S.W.; Tsui, K.-L.; Ariyajunya, B.; and Kim, S.B. (2009). "A Comparison of CUSUM, EWMA, and Temporal Scan Statistics for Detection of Increases in Poisson Rates." *Quality and Reliability Engineering International* 26, pp. 279-289.

Lucas, J.M. (1985). "Counted data CUSUM's." *Technometrics* 27(2), pp. 129-144.

Montgomery, D. C. (2008). *Introduction to Statistical Quality Control*. Sixth Edition. Hoboken, N.J: John Wiley.

Occupational Safety and Health Administration (OSHA) (2013). "Commonly Used Statistics." Retrieved from <http://www.osha.gov/oshstats/commonstats.html>

Page, E.S. (1954). "Continuous Inspection Schemes." *Biometrika* 41, pp. 100-115.

Reynolds, M.R., Jr., and Stoumbos, Z.G. (1999). "A CUSUM Chart for Monitoring a Proportion When Inspecting Continuously." *Journal of Quality Technology* 31(1), pp. 87-108.

Reynolds, M.R., Jr., and Stoumbos, Z.G. (2000). "A General Approach to Modeling CUSUM Charts for a Proportion." *IIE Transactions* 32, pp. 515-535.

Reynolds, M.R., Jr., and Stoumbos, Z.G. (2004a). "Control Charts and the Efficient Allocation of Sampling Resources." *Technometrics* 46(2), pp. 200-214.

Reynolds, M.R., Jr., and Stoumbos, Z.G. (2004b). "Should Observations Be Grouped for Effective Process Monitoring?" *Journal of Quality Technology* 36(4), pp. 343-366.

Schuh, A., Camelio, J.A., and Woodall, W.H. (2012). "Control Charts for Accident Frequency: A Motivation for Real-Time Occupational Safety Monitoring." Accepted for publication in *International Journal of Injury Control and Safety Promotion*.

Steiner, S. H. and Jones, M. (2010). "Risk-Adjusted Survival Time Monitoring with an Updating Exponentially Weighted Moving Average (EWMA) Control Chart." *Statistics in Medicine* 29, pp. 444-454.

Szarka, J.L., III and Woodall, W.H. (2011). "A Review and Perspective on Surveillance of Bernoulli Processes." *Quality and Reliability Engineering International* 27, pp. 735-752.

Szarka, J.L., III and Woodall, W.H. (2012). "On the Equivalence of the Bernoulli and Geometric CUSUM Charts." *Journal of Quality Technology* 44(1), pp. 54-62.

White, C.H.; Keats, J.B.; and Stanley, J. (1997). "Poisson CUSUM Versus c Chart for Defect Data." *Quality Engineering* 9(4), pp. 673-679.

Wu, Z.; Liu, Y.; Zhen, H.; and Khoo, M. (2010). "A Cumulative Sum Scheme for Monitoring Frequency and Size of an Event." *Quality and Reliability Engineering International* 26, pp. 541-554.

The following manuscript, authored by Anna Schuh and Jaime A. Camelio, is copyrighted by the Institute of Industrial Engineers to appear in the *2013 Proceedings of the Industrial and Systems Engineering Research Conference* in San Juan, Puerto Rico, May 18-22, 2013. Reprinted with the permission of the Institute of Industrial Engineers, 3577 Parkway Lane, Suite 200, Norcross, GA 30092, www.iienet.org Copyright©2012 by Institute of Industrial Engineers.

5. Including Accident Severity in Statistical Monitoring Systems for Occupational Safety

5.1 Abstract

In order to reduce the rates of occupational fatalities and nonfatal injuries in industry, it is important to measure, monitor, and report safety performance over time, and act accordingly upon any findings. Real-time statistical monitoring for occupational safety has focused mostly on monitoring and improving upon accident frequency considerations. However, another important aspect of occupational safety incidents is their severity. If safety initiatives need to be prioritized, obviously those that address and reduce the most severe accidents should be implemented first. Severity has been considered in safety improvement processes before, such as the Risk Assessment Matrix or the Department of Energy Cost Index, but it is not clear how these definitions of severity could be transferred to quantitative statistical monitoring systems. This paper reviews current accident severity metrics, including the monitoring of OSHA-recordable “Days Away from Work” and “Days in Job Restriction/Transfer.” Recommendations for the selection of a severity metric and a temporal directionality of the metric (reactive, predictive, or proactive) for statistical monitoring of accident severity are given.

5.2 Introduction

The United Nation’s International Labour Organization (ILO) estimates that there are annually 337 million occupational accidents and 160 million occupational diseases, resulting in a total of 2.3 million worker deaths per year (Niu, 2010). In the U.S., approximately 4,600 fatal work injuries occurred in 2011 (Bureau of Labor Statistics, BLS, 2013). In an effort to reduce and eliminate these accidents, the National Institute of Occupational Safety and Health (NIOSH) has made surveillance a priority in their strategic goals (NIOSH, 2011). Currently, the Occupational Safety and Health Administration (OSHA) collects required reporting materials from companies employing more than 11 workers at the end of each calendar year, and the Bureau of Labor Statistics (BLS) releases this information to the public in September of the following year (BLS,

2013). This obvious data lag may trickle down to companies and industries who may not collect or analyze data on their own, outside of OSHA requirements. This causes a similar lag in root-cause investigation which usually occurs on the smaller scale. It has been suggested by many authors that more frequent monitoring of accidents using statistical monitoring can improve the ability to signal unsafe conditions (Wu et al., 2010; Lucas, 1985; Schuh et al., 2013a; Schuh et al., 2013b; Schuh et al., 2013c).

Applications of real-time statistical monitoring for safety have focused on only monitoring accident frequency (e.g., Blindauer and Michael, 1959; Adekeye and Aluko, 2012; Cournoyer et al., 2011), including initial attempts for specifically monitoring industrial accidents (Lucas, 1985; Schuh et al., 2013a; Schuh et al., 2013b; Schuh et al., 2013c). However, accident severity is also important for prioritizing the implementation of new best practices. A core challenge to injury surveillance is defining and agreeing upon what constitutes an “important injury event,” (Langley, 2004) and investigating relative severity may be one way to accomplish this.

This paper introduces the challenges associated with the active monitoring of accident severity. Section 5.3 discusses how statistical surveillance has recently been used to monitor accident frequency. In Section 5.4, the selection or development of a severity metric is encouraged and some common accident severity metrics are reviewed. Section 5.5 motivates the choice of a safety management approach for monitoring performance indicators, and outlines how reactive, predictive and proactive indicators could be used in the statistical monitoring of accident severity metrics. Finally, conclusions and suggestions for future work are offered in Section 5.6.

5.3 Statistical Monitoring Background

A statistical monitoring tool that has traditionally been used for manufacturing applications is the statistical process control chart. In the 1920s, Walter Shewhart introduced these charts in order to monitor the number of defective parts and address machining problems as soon as possible when an unusual number of defectives was observed (Montgomery, 2008).

More recently, statistical process control charts have been adapted for use in healthcare applications (Woodall, 2006; Woodall et al., 2010). When specifically considering accident frequency, initial studies have found that monitoring real-time data with an exponential cumulative sum (CUSUM) control chart will signal safety problems sooner than using a Poisson

CUSUM chart to monitor accident data that has been aggregated over a specific period of time (e.g., a year, a quarter, or a month) (Schuh et al., 2013b; Schuh et al., 2013c). An example of an exponential CUSUM control chart that could be used for monitoring accidents in real time is given in Figure 1, where a signal is given after about 61 days of monitoring (Schuh et al., 2013b).

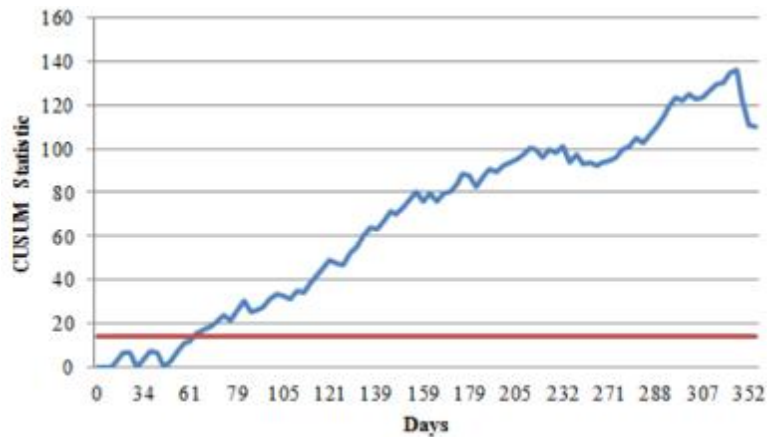


Figure 5.1: Exponential CUSUM control chart for monitoring accident frequency (from Schuh et al., 2013b)

The purpose of using a similar control chart system for monitoring accident severity would be to prioritize the implementation of new best practices. For instance, when monitoring accident frequency alone, it might be difficult for safety officials to immediately recognize which accident types are resulting in the most severe outcomes. If companies and industries were tracking an accident severity metric alongside accident frequency, hazard controls could more easily be implemented in order of urgency, according to the degree of severity associated with the accident they are intended to prevent. The monitoring of accident severity with a control chart has only been briefly suggested for the small subset of accidents resulting in multiple fatalities (Wu et al., 2010).

In order to develop a system for monitoring the severity of all accidents in real-time, a current severity performance metric will need to be chosen (or a new one created), and a monitoring directionality will need to be selected (backward, present, or forward). Section 5.4 reviews the benefits and limitations of some common severity metrics that are currently used, and discusses their shortcomings for use in real-time or near-real-time statistical monitoring systems. In Section 5.5 the directionality of monitoring metrics, including reactive, predictive and proactive

implementation methods are discussed. While these topics are reviewed in depth here, it will be imperative for future studies to investigate the various combinations of metrics and monitoring methods.

5.4 Severity Metrics

First, it will be necessary to choose a severity metric to monitor. Many metrics for accident severity are currently being used. This section reviews these metrics and discusses their benefits and limitations for use in active statistical monitoring systems.

5.4.1 Risk Assessment Matrix

The Risk Assessment Matrix is a tool used to determine an activity’s risk based on the probability and severity of the potential accidents associated with the activity. This matrix was originally developed for general safety analysis, but has been adapted for use in the U.S. military. The version adapted by the U.S. Department of Defense is given in Figure 5.2.

RISK ASSESSMENT MATRIX				
SEVERITY \ PROBABILITY	Catastrophic (1)	Critical (2)	Marginal (3)	Negligible (4)
Frequent (A)	High	High	Serious	Medium
Probable (B)	High	High	Serious	Medium
Occasional (C)	High	Serious	Medium	Low
Remote (D)	Serious	Medium	Medium	Low
Improbable (E)	Medium	Medium	Medium	Low
Eliminated (F)	Eliminated			

Figure 5.2: Risk Assessment Matrix (from U.S. Department of Defense, 2012)

The risk assessment matrix has been criticized for being too subjective and not quantitative enough to effectively mitigate risk (Cox, 2008). Also, actively monitoring changes in these risk levels will be difficult since there is no associated quantitative value. Therefore, monitoring these risk level may not be the most effective way to incorporate incident severity into statistical safety monitoring.

5.4.2 Lost and Restricted Workday Metrics

Currently, OSHA and BLS recognize the lost workday (LWD) rate as a key measure of workplace injury and illness severity (BLS, 2012). This measure is calculated as the ratio of the number of employee days away from work and the total number of hours worked in a firm during a year. The effect of this measure has been analyzed in past studies. For instance, the historically most common accident types, injury sources, and body parts associated with accidents resulting in 21 or more days away from work has been studied (Mital et al., 1999), lost work time and workers compensation metrics in the construction industry have been used to identify areas where more safety resources are needed (Lowery et al., 2000), and an Injury Severity Score has been developed, which combines the days away from work with the number of days spent in the hospital, as well as indices for how maimed or fractured an employee was (Larsson and Field, 2002). An extensive literature review summarizing the historical uses of the lost work time metrics in the mining industry has also been conducted (Coleman and Kerkerling, 2007). The DART rate (Days Away, Restricted or Transferred) is a similar metric that has recently replaced the LWD metric for OSHA and BLS severity assessments.

One obvious shortcoming of these rates is that they are clearly lagging indicators. That is, they only assess past incidents. “Leading” indicators, on the other hand, would instead attempt to predict and prevent future accidents without necessarily relying on past data (Hinze et al., 2013). However, the LWD and DART rates are both products of the OSHA and BLS annual reports, and it would be difficult to measure and monitor them much more frequently. For example, if an injury is classified as “severe” based on the number of days away from work (and/or in job restriction or transfer), an employee would need to be out of work (or restricted or transferred) for a long period of time in order to be considered “severe.” If one were to attempt to monitor a severity metric monthly, for example, it is possible that some employees will not have completed their time away/restricted by the end of a given month, and therefore the number of days away/restricted would not accurately represent the accident’s “severity.” The conundrum of this updating metric is shown in Figure 5.3.

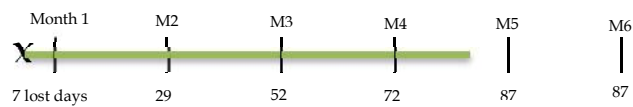


Figure 5.3: Total lost work days resulting from a safety incident, count updated monthly

In order to monitor a severity metric in real time, it may be desirable to predict the LWD or DART rates from the onset of an injury. However, it has been found that emergency physicians showed very low accuracy when predicting the length of patients' temporary disability (Beach et al., 2012). Several previous studies have been conducted to predict the length a of disability, and it has been found that many factors can influence this highly variable value, including age, weight, gender, experience level, socioeconomic status, and many other details [e.g., Cheadle et al., 1994; Krause et al., 2001; MacEachon et al., 2006). Therefore, it may be difficult to accurately predict this metric for real-time monitoring.

5.4.3 Department of Energy Severity Cost Index

The Department of Energy has developed a Safety Cost Index, defining the severity of an occupational accident by its expected associated monetary cost (Gochfield and Mohr, 2007). This cost equation is given by Equation 1, where C represents the approximate cost of the incident:

$$C = 100 * \frac{(1,000,000 * D) + (500,000 * T) + (2,000 * LWC) + (1,000 * WDL) + (400 * WDLR) + (2,000 * NFC)}{\text{Total Hours Worked}} \quad (1)$$

The variables used in Equation 1 are defined as:

Accident Category	Weight	Equation Symbol
# of deaths	1,000,000	D
# of permanent job transfers, restrictions or terminations	500,000	T
# of cases with lost or restricted workdays	2,000	LWC
# of cases with no lost time	2,000	NFC
# of lost days	1,000	WDL
# of restricted days	400	WDLR

The Computerized Accident/Incident Reporting System (CAIRS) is a statistical monitoring tool that has been developed based on this cost index (Kenoyer et al., 2001).

This performance metric is easily utilized with yearly aggregated reporting methods, but challenges to its use are seen when monitoring is implemented over shorter time periods. Since the number of days away or restricted is often not known until long after the accident occurs, this

metric - like the LWD or DART rates - cannot be accurately calculated and actively monitored in real-time or near real-time.

5.4.4 Survival analysis

Based on the drawbacks presented, it can be seen that it is difficult to monitor any of these three accident severity metrics in real time. A potential statistical monitoring method for considering a length of time in a given state as it accrues is through “survival analysis” of “lifetime data,” as presented by many authors. This analysis monitors the length of time until a defined success or failure. For example, it has been used to monitor the length of time a patient survives following surgery (Steiner and Jones, 2009), the lifetimes of computer component parts (Yaschin, 2010), and unemployment durations (Kiefer, 1988). The data in these studies have often not reached the defined success or failure (death, component failure, or employment) at the time of data collection, and are therefore referred to as “censored.” Survival analysis has been extended to traffic safety for considering the lengths of road trips before an automobile accident occurs (Chang and Jovanis, 1990).

However, this survival analysis generally only reports the time until the given failure or success. That is, for the computer component example, the data collection only reports whether the computer component has failed or is still functioning at a given time. Unfortunately, since our goal is to actively monitor accident severity, it would not be sufficient simply to know whether an employee remains disabled. Rather, it would be more beneficial to know how many days the employee has been out-of-work or in restriction at each monitoring point. For instance, if a company were to monitor on a weekly basis, they would prefer to know at the end of each week how many total days the employee has been out of work. Since this variable would update every week, there would be many complications to monitoring this changing variable that have not been previously investigated for survival analysis. This challenge will be especially prevalent when monitoring in real-time instead of in aggregated periods.

In order to develop a monitoring system for accident severity, it is essential that one of these current performance metrics be used, or that a new one be developed. Once a severity metric has been chosen, safety officials can choose an appropriate reactive, predictive or proactive

monitoring method. Section 5.5 will review these monitoring strategies and discuss how they would be employed specifically for statistical monitoring of accident severity.

5.5 Safety Management System Methods for Accident Severity

Safety management systems have often employed systematic safety performance strategies for temporal directionality, which involve reactive, predictive, and proactive approaches (International Civil Aviation Organization, 2009; Hopkins, 2009; Vredenburg, 2002; K rvers and Sonnemans, 2008). These strategies will be adapted for a discussion of statistical monitoring of accident severity. Since the benefits and limitations of each of these approaches will be highly dependent upon the chosen metric, the authors leave it to the practitioner to investigate the most appropriate method for each particular metric.

5.5.1 Reactive

A reactive approach would involve intensive root-cause investigations of past severe accidents. The inherent assumption of choosing this approach is that the causes of past accidents will be the same as future accidents. Under this assumption, the accurate identification of past accident sources can lead to new hazard controls that will eliminate the causes and prevent similar accidents in the future. When monitoring severe accidents with a control chart, this reactive approach will involve an analysis of the change point, or the point in time during which the control chart began to shift out of control. This will be earlier than the process signal, and will usually be the point at which a problem occurred. Therefore, identifying this point is critical to the success of the reactive approach. This change point is found using a bootstrapping procedure to determine confidence levels for CUSUM control charts, and this analysis has been conducted in the past for various applications (Kass-Hout et al., 20012; Shao and Hou, 2012; Mahmoud, 2007). A control chart signal for the reactive monitoring of a severity metric would indicate the need to investigate the change point.

One obstacle to root-cause change-point investigation is the potential data lag that can occur when monitoring accident severity. As mentioned in Section 5.4, several current severity metrics won't classify an incident as a certain degree of "severe" until a significant amount of time has passed (e.g., a certain number of lost workdays have occurred or a specific monetary loss has accumulated). By the time the incident is classified as severe and the need for investigation is

established through statistical monitoring and other tools, it may be difficult to accurately perform root-cause analysis.

Another shortcoming of the reactive method lies in the inherent assumption that past accident causes will be the same as future ones. Especially in a dynamic industrial environment where new projects, technologies, and methodologies are continually being adopted, old accident causes may become obsolete as new ones are introduced and it will therefore not be sufficient to only react to past incidents.

5.5.2 Severity Predictive

The predictive approach that is usually included in discussions of event management involves hazard prediction, which is presented in Section 5.5.4. Since severity is not typically a part of these considerations, we have included a new severity predictive strategy, which is the prediction of an accident's severity immediately following the incident. This approach addresses the data lag problem seen in the reactive method. Based on a variety of immediately known key performance indicators, a prediction of the eventual value of the severity metric in real time would also be plotted as soon as possible after the accident occurs, and a signal would indicate an increase in the number of accidents that are expected to be designated as severe.

Unfortunately, not all metrics will be easily and accurately predicted. For example, as discussed in Section 5.4, the number of lost workdays resulting from an accident is difficult for even a physician to predict (Beach et al., 2012). One solution to this problem might be to instead use a classification algorithm to assign a predicted severity score to an accident at its onset. This scoring concept is already utilized with the Injury Severity Score to assign a score long after the incident by combining the number of lost workdays with the length of hospitalization, how maimed the employee was, and how fractured they were. These first two measures in this score require a data lag which would hinder the predictive approach, but the latter two measures are factors which can be assessed relatively quickly following an accident. Therefore, for this severity predictive approach, a scoring system or classification algorithm should only include these types of easily determined factors. It should be noted that a score or classification scale will produce a discrete variable, which will affect the type and use of control charting systems.

5.5.3 Proactive

Rather than consider any past information, proactive methods will attempt to eliminate hazards associated with severe accidents before the severe accidents actually occur. A common proactive safety method is to record near misses, or near hits, which are incidents which do not lead to recordable injuries, but have that potential. A review of near misses and similar leading indicators is given by (Hinze et al., 2013). A control chart signal when monitoring near misses would indicate an increase in potentially unsafe conditions.

Monitoring near misses would provide valuable front-end information that could potentially correct hazards before disastrous accidents occur, but the metric has been widely criticized for under-reporting (Hinze et al., 2013). Unfortunately, the benefits of statistical monitoring are greatly lessened when there is an unreliable data stream, so this strategy should be accompanied by efforts to encourage employee reporting of these proactive indicators.

5.5.4 Hazard Predictive

The last event management strategy involves the prediction of future safety hazards. This method seeks to conclude sources of potential future risk as new practices and technologies are implemented, and attempt to mitigate the sources of these potential risks before they become hazards. An example of a hazard predictive metric is the safety climate, or employees' perception and attitude regarding the safety environment (Clarke, 2006). If employees are concerned about their workplace, there may be hidden risks that have not been previously identified. When monitoring the changes in these attitudes over time, a control chart signal would indicate a potential for future risks.

However, these predictive metrics have been criticized for not strongly predicting future accidents since perceptions and attitudes are highly unique to an individual (e.g., Clarke, 2006; Nielsen and Mikkelsen, 2007; Henning et al., 2009). Therefore, these metrics may require further investigation before this strategy is relied upon. Since this strategy has the highest potential to curtail future accidents, future work should consider the development of hazard predictive metrics that are both more reliable and can be monitored in real-time with control charts.

The differences between the four safety management approaches are summarized in Table 5.1.

Approach	Focus	Control Chart Signal Indication	Limitations
Reactive	Past accidents	A need to investigate the change point	Data lag
Severity Predictive	Current accidents	An increase in accidents that are expected to be classified as “severe”	Many severity metrics are not easily predicted
Proactive	Current hazards	An increase in unsafe conditions	Leading indicators like near misses are often underreported
Hazard Predictive	Future hazards	A potential for future hazards	Current indicators may not be strong predictors

Table 5.1: Safety Management System approaches applied to control charting of severe accidents

If a quantitative severity metric was monitored in real time with a CUSUM control chart, Figure 5.4 indicates the points of interest for each of the monitoring strategies. Note that the CUSUM statistic is reset to zero following a signal, since it is expected that, in practice, safety officials will implement new safety initiatives following a signal and the CUSUM value will be reset, as in Schuh et al., 2013b.

Again, while there may be benefits and limitations of each of these directionalities, selecting the “best” method for a particular monitoring system is left up to the safety officials familiar with that particular system.

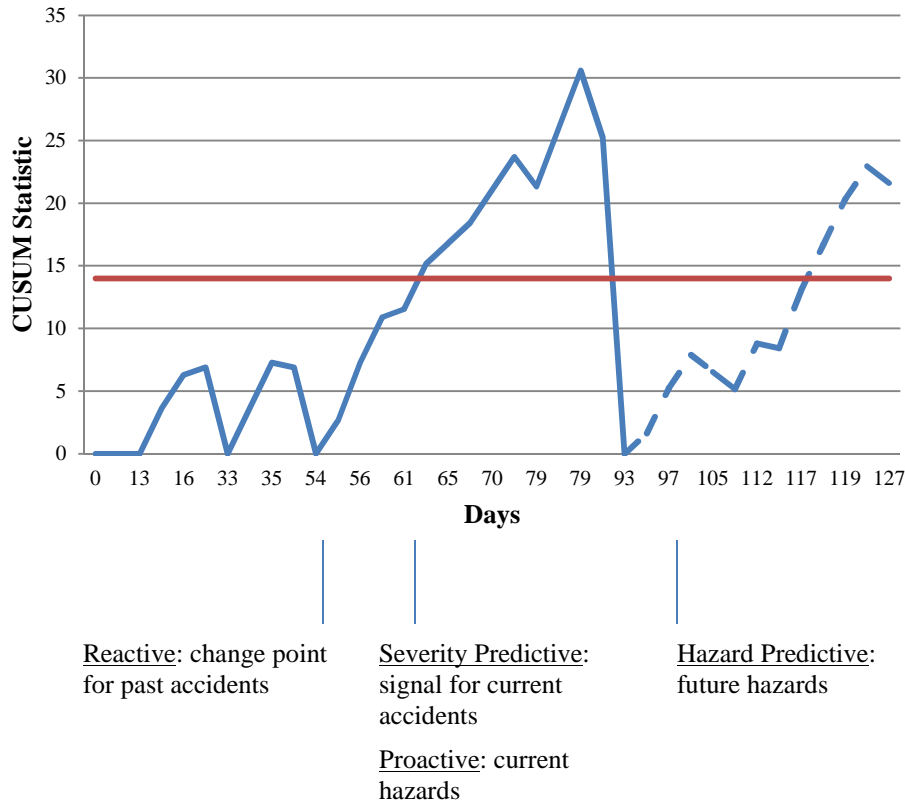


Figure 5.4: Event management points of interest with a control charting system

5.6 Conclusions and Future Work

This work has motivated the need to extend current statistical monitoring methods for accident frequency to include accident severity. Two necessary tools for this endeavor will include the choice of a severity metric and an appropriate monitoring directionality for that metric. The common severity metrics that are currently being used were reviewed and the challenges associated with monitoring their performance using statistical process control charts were discussed. Common safety management strategies for monitoring data with reactive, predictive, and proactive directionalities were introduced and their specific future application to control chart monitoring of accident severity was considered.

Real-time monitoring of accident severity has rarely been investigated in the past applications. Therefore, this initial motivational discussion is meant to encourage extensive future work.

5.7 References

- Adekeye, K.S., and Aluko, O.S., 2012, "Design of Cusum Scheme for Monitoring Road Accident Fatalities," *Open Journal of Statistics*, 2, 213-218.
- Beach, J., Benoit, M., Rowe, B.H., and Cherry, N, 2012, "Can Emergency Physicians Predict Severity and Time Away from Work?" *Occupational Medicine*.
- Blindauer, H.H. and Michael, H.L., 1959, "An Analysis of High Accident Rates: Technical Paper." Publication FHWA/IN/JHRP-59/14. Joint Highway Research Project, Indiana Department of Transportation and Purdue University, West Lafayette, IN.
- Bureau of Labor Statistics (BLS) (2013). "Injuries, Illnesses and Fatalities." Retrieved from <http://bls.gov/iif/>
- Centers for Disease Control and Prevention (CDC), National Institute of Occupational Safety and Health (NIOSH), 2013, "NIOSH Strategic Goals: Surveillance," <http://www.cdc.gov/niosh/programs/surv/goals.html>
- Chang, H.-L. and Jovanis, P.P, 1990, "Formulating Accident Occurrence as a Survival Process," *Accident Analysis and Prevention*, 22(5), 407-419.
- Cheadle, A., Franklin, G., Wolfhagen, C., Savarino, J., Liu, P.Y., Salley, C., and Weaver, M., 1994, "Factors Influencing the Duration of Work-Related Disability: A Population-Based Study of Washington State Workers' Compensation," *American Journal of Public Health*, 84(2), 190-196.
- Clarke, S., 2006, "Contrasting Perceptual, Attitudinal and Dispositional Approaches to Accident Involvement in the Workplace," *Safety Science*, 44, 537-550.
- Coleman, P.J. and Kerkering, J.C., 2007, "Measuring Mining Safety with Injury Statistics: Lost Workdays as Indicators of Risk," *Journal of Safety Research*, 38, 523-533.
- Cournoyer, M., Garcia, V., Gallegos, U., and Wilburn, D., 2011, "Investigation of Injury/Illness Data at a Nuclear Facility," *Journal of Chemical Health and Safety*, 8(5), 17-25.
- Cox, L.A., Jr., 2008, "What's Wrong With Risk Matrices?" *Risk Analysis*, 28(2), 498-512.
- Gochfield, M. and Mohr, S., 2007, "Protecting Contract Workers: Case Study of the U.S. Department of Energy's Nuclear and Chemical Waste Management," *American Journal of Public Health*, 97(9), 1607-1613.
- Henning, J.B., Stufft, C.J., Payne, S.C., Bergman, M.E., Mannan, M.S., and Keren, N., 2009, "The Influence of Individual Differences on Organizational Safety Attitudes," *Safety Science*, 47, 337-345.

Hinze, J., Thurman, S., and Wehle, A., 2013, "Leading Indicators of Safety Performance," *Safety Science*, 51, 23-28.

Hopkins, A., 2009, "Thinking About Process Safety Indicators," *Safety Science*, 47, 460-465.

International Civil Aviation Organization, 2009, "Safety Management Manual," Doc 9859.

Kass-Hout, T.A., Xu, Z., McMurray, P., Park, S., Buckeridge, D.L., Brownstein, J.S., Finelli, L., and Groseclose, S.L., 2012, "Application of Change-Point Analysis to Daily Influenza-Like Illness Emergency Department Visits," *Journal of the American Medical Informatics Association*, 19, 1075-1081.

Kenoyer, J., Andrews, W., and Stenner, R., 2001, "Monitoring and Predicting Worker Risk," *Environmental Science and Pollution Research International*, 8(1), 36-44.

Kiefer, N. M., 1988, "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26(2), 646-679.

Krivers, P.M.W. and Sonnemans, P.J.M, 2008, "Accidents: A Discrepancy Between Indicators and Facts!" *Safety Science*, 46, 1067-1077.

Krause, N., Frank, J.W., Dasinger, L.K., Sullivan, T.J., and Sinclair, S.J., 2001, "Determinants of Duration of Disability and Return-to-Work After Work-Related Injury and Illness: Challenges for Future Research," *American Journal of Industrial Medicine*, 40, 464-484.

Langley, J., 2004, "Challenges for Surveillance for Injury Prevention," *International Journal of Injury Control and Safety Promotion*, 11(1), 3-8.

Larsson, T.J. and Field, B., 2002, "The Distribution of Occupational Injury Risks in the Victorian Construction Industry," *Safety Science*, 40, 439-456.

Lowery, J.T., Glazner, J., Borgerding, J.A., Bondy, J., Lozotte, D.C., and Kreiss, K., 2000, "Analysis of Construction Injury Burden by Type of Work," *American Journal of Industrial Medicine*, 37, 390-399.

Lucas, J.M., 1985, "Counted data CUSUM's," *Technometrics*, 27(2), 129-144.

MacEachon, E., Clarke, J., Franche, R.-L., and Irvin, E., 2006, "Systematic Review of the Literature on Return to Work After Injury," *Scandinavian Journal of Work, Environment and Health*, 32(4), 257-269.

Mahmoud, M.A., Parker, P.A., Woodall, W.H., and Hawkins, D.M., 2007, "A Change Point Method for Linear Profile Data," *Quality and Reliability Engineering International*, 23, 247-268.

- Mital, A., Pennathur, A., and Kansal, A., 1999, "Nonfatal Occupational Injuries in the United States Part I – Overall Trends in Data Summaries," *International Journal of Industrial Ergonomics*, 25, 109-129.
- Montgomery, D. C., 2008, *Introduction to Statistical Quality Control*. Sixth Edition. Hoboken, N.J: John Wiley.
- Nielsen, K.J. and Mikkelsen, K.L., 2007, "Predictive Factors for Self-Reported Occupational Injuries at 3 Manufacturing Plants," *Safety Science Monitor*, 2 (11), 1-9.
- Niu, S., 2010, "Ergonomics and Occupational Safety and Health," *Applied Ergonomics*, 41, 744-753.
- Schuh, A., Camelio, J.A., Kleiner, B.M., 2013a, "Traditional Performance Measurement in Construction Safety: A Case for Active Safety Monitoring," Submitted for publication.
- Schuh, A., Camelio, J.A., and Woodall, W.H., 2013b, "Control Charts for Accident Frequency: A Motivation for Real-Time Occupational Safety Monitoring," Accepted for publication in *International Journal of Injury Control and Safety Promotion*.
- Schuh, A., Woodall, W.H., and Camelio, J.A., 2013c, "The Effect of Aggregating Data When Monitoring a Poisson Process," Accepted for publication in *Journal of Quality Technology*.
- Shao, Y.E. and Hou, C.-D., 2012, "Change Point Determination for a Multivariate Process Using a Two-Stage Hybrid Scheme," *Applied Soft Computing*.
- Steiner, S.H. and Jones, M., 2009, "Risk-Adjusted Survival Time Monitoring with an Updating Exponentially Weighted Moving Average (EWMA) Control Chart," *Statistics in Medicine*, 29, 444-454.
- U.S. Department of Defense MIL-STD-882E, 2012, "Department of Defense Standard Practice for System Safety."
- Vredenburgh, A.G., 2002, "Organizational Safety: Which Management Practices Are Most Effective in Reducing Employee Injury Rates?" *Journal of Safety Research*, 33, 259-276.
- Woodall, W.H., 2006, "The Use of Control Charts in Health-Care and Public-Health Surveillance," *Journal of Quality Technology*, 38(2), 89-134.
- Woodall, W.H., Grigg, O.A., and Burkom, H.S., 2010, "Research issues and ideas on health related surveillance," *Frontiers in Statistical Quality Control*, 9, edited by H.J. Lenz, P.-Th. Wilrich, and W. Schmid, Berlin-Heidelberg, Germany: Springer-Berlag, 145-155.
- Wu, Z., Liu, Y., Zhen, H., and Khoo, M., 2010, "A Cumulative Sum Scheme for Monitoring Frequency and Size of an Event," *Quality and Reliability Engineering International*, 26, 541-554.

Yashchin, E., 2010, “Computational and Monte-Carlo Aspects of Systems for Monitoring Reliability Data,” IBM Research Report RC25029

6. Conclusions and Future Work

Motivated by the limited public safety data available, the overall conclusion of this work is that smaller periods of aggregation are preferred. Whether monitoring Safety Incident Indicator values, CUSUM control charts, or an accident severity indicator, this dissertation work has shown that it is desirable to consider updating and reporting public health and safety data frequently with the lowest possible degree of aggregation. The presented work has demonstrated the importance of reducing aggregation windows so safety problems can be detected in a timely manner and practices can be put in place to mitigate future incidents. The effectiveness of these newly implemented hazard controls can also be tracked using these monitoring systems, but a timely data stream will be required for all of these potential benefits to be effectively realized. The contributions provided by this dissertation are summarized in Section 6.1.

The next sections introduce areas for future work, based on the limitations that were discovered throughout the presented research. Many of the limitations cited in the four presented manuscripts can be addressed with better data collection practices, so a detailed summary of desired data fields is given in Section 6.2. Section 6.3 discusses some extensions for future control chart monitoring systems in the safety domain. Finally, Section 6.4 discusses how the fields of occupational safety and statistical monitoring can continue to advance together through good communication processes and by ensuring the usability of all implemented systems.

6.1 Contributions

This dissertation has used statistical surveillance methods to show that shorter time periods of aggregation are desired when monitoring occupational safety data. Less aggregation will allow for timely implementation of hazard controls, which will hopefully reduce the frequency of accidents and save lives. The effects of data aggregation for process monitoring have been considered in the past, but very minimally for the specific application to public health and occupational safety surveillance.

First, the Hierarchy of Hazard Controls was used to motivate the need for more timely accident investigations, in order to implement the most effective hazard controls possible. A previous program for quality improvement in the automobile industry was used as a framework to design

a new method to specifically monitor and report the most frequent accident types. By focusing safety efforts on these most problematic areas, it was suggested through a case study that more frequent monitoring will reduce accidents.

Chapter 3 presented a more quantitative approach and considered the use of statistical process control charts for the monitoring accident frequency. Under the assumption of a Poisson process, the time-to-signal performance of Poisson CUSUM control charts with data collected in various aggregation periods was compared to that of exponential CUSUM control charts with non-aggregated data, building upon limited past comparisons of these charts. For one example of real data, it was shown that exponential CUSUM charts have slightly better performance than daily Poisson CUSUM charts. The benefits of the exponential charts improve when shorter hazard control implementation times are assumed and non-zero restarting parameters are used.

Next, a more general comparison of Poisson and exponential CUSUM control chart performance was presented using simulated data. Expanding one earlier work in which zero-state exponential CUSUM charts were compared to Poisson CUSUM charts for short periods of aggregation, steady state charts with longer aggregation periods were considered in order to specifically consider the results for public health and safety applications. It was shown that the exponential CUSUM chart for non-aggregated data had the best ATS detection performance. Furthermore, it was established that many additional adverse events (accidents) will occur between the control chart signal from an exponential chart and that from a long-term aggregated Poisson chart.

Finally, a brief discussion about the future consideration of accident severity in statistical monitoring systems was presented in Chapter 5. Current severity metrics were reviewed and incident management strategies for reactive, predictive and proactive methodologies were discussed. Extensive future work in this area is needed and encouraged.

This work has concluded that for occupational safety surveillance, incidents should be reported as often as possible and monitored with the smallest possible levels of aggregation. These conclusions advance work in both the occupational safety and statistical monitoring fields. Possible topics for future work are suggested in the following sections.

6.2 Data Collection and Analysis

Throughout this work, a lack of available information has been consistently noted as a significant limitation. The data fields that would be desired to successfully implement the safety monitoring and surveillance systems presented in the preceding chapters, and advance future work in safety monitoring systems, are discussed below.

6.2.1 More Frequent Data Collection

This has been the comprehensive outcome of the conducted research reported in Chapters 2-5. When data is collected frequently and actively monitored, problems can be signaled at their onset and new hazard controls for accident mitigation can be introduced. The effectiveness of newly implemented hazard controls can be confirmed in a more timely fashion as well.

6.2.2 Employee and Environment Data

Future data collection strategies should focus on the goal of surveillance systems and therefore include many descriptors of safety incidents. For instance, environmental circumstances could influence the likelihood of accidents. Seasonal effects could be present in some industries (Leamon and Murphy, 1995; Lipscomb et al., 2006), or accident occurrence may be more likely on certain days of the week (Brogmus, 2007; Wirtz, 2011). While OSHA requires the date of an accident to be recorded, analysis of these data trends may not be performed regularly. Likewise, individual employees may be more at risk for certain types of incidents; age, overall mental and physical health, degree of obesity, and level of experience could all influence an employee's likelihood for accident occurrence. Collecting this data and monitoring trends over time should lead to more informed training courses for higher-risk employee subgroups and better prevention strategies for common environmental contributions. There are many factors that influence workplace safety (Subramanian, 2006), so safety officers will need to work closely with statisticians to accurately define the needed data categories and appropriate quantitative metrics for each piece of data. Risk adjustment for these factors in control charting monitoring systems is discussed in Section 6.3.

6.2.3 Leading Indicators

It has been established in the earlier chapters that we can monitor and signal process changes, and that frequent data collection will allow for timely reactive responses. However, a proactive approach in which safety officials receive data about current hazards would have the potential to indicate an increase in hazards before any recordable injuries occur as a result of them. This would obviously be preferable, but sufficient leading indicator data would need to be collected in order for this type of monitoring system to be reliable.

6.3 Control Charting Extensions

Chapters 3 and 4 of this dissertation considered the effects of data aggregation on the ATS performance of Poisson and exponential CUSUM control charts. This analysis provided an initial discussion about data aggregation when monitoring public health data, but extensions to this work are still necessary. Specifically, some of these extensions might include the consideration of risk adjustment, transient shifts, control charts for monitoring process improvements instead of only deterioration, the presence of autocorrelated data, and the use of other distributions to describe the time between events. These additional considerations are specifically necessary for extensions to the monitoring systems for accident frequency discussed in Chapter 3 and 4; of course even more considerations will be necessary if severity is also included, as briefly discussed in Chapter 5.

6.3.1 Risk Adjustment

Risk adjustment was first introduced for use in statistical healthcare surveillance by Lovegrove et al. (1997, 1999) and Poloniecki et al. (1998). Steiner et al. (2000) monitored surgical deaths with CUSUM charts, using risk adjustment to place a higher weight on the death of someone with a low probability of death upon entering surgery. The Parsonnet scoring system was used to determine the pre-operative likelihood of surgical death for each patient based on a series of personal attributes, including age, weight, and health. In the situation that a surgical death occurred with a patient who was determined to have a low probability of death, it is more likely that an assignable human error was the cause of this death than in the case of a patient with higher probability of death at the onset of surgery. In this way, risk adjustment encourages the investigation of preventable errors during surgery, so that future deaths can be mitigated.

Extensions of this application have been explored at length (e.g., Steiner et al., 2001; Grigg and Farewell, 2004; Segó et al., 2009).

The application of this principle should be considered when collecting occupational safety data. Similar to the uses in healthcare, an employee's susceptibility to accidents would be quantified based on various personal attributes and incidents involving those with lower risk would be weighted more significantly in a monitoring system. Unlike healthcare, risk assessment for safety could also include environmental factors. For instance, accidents occurring in less likely seasons, on less likely days, or in less risky operations, would be more heavily weighted in a risk-adjusted statistical monitoring system for safety. As noted in Section 6.2, collection of the necessary data will be essential for the success of risk adjustment methods.

It should be noted that just because an accident is more probable based on any of these factors does not indicate that the occurrence of the incident is acceptable; appropriate efforts should be made to reduce and eliminate all accidents. Again, this risk adjustment approach is meant to prioritize the root-cause investigation of assignable errors based on an incidents' relative probability of occurrence.

6.3.2 Transient Shifts

Chapter 4 of this dissertation considers the effect of data aggregation for sustained process shifts. However, it is possible that public health and safety data may experience transient shifts, where the process briefly shifts to out-of-control behavior and then naturally returns to in-control behavior without any changes to the process. In this case, the process shift will likely not be noticed if data is being aggregated over long periods of time and the shift is not long enough to significantly affect the process mean count of accidents. The root-cause of the shift would not be investigated and would likely continue to occur unnoticed, until the problem progressed. It should be noted that the return to in-control behavior could be the result of successful hazard controls that are already implemented on a worksite, which is a promising indicator of their effectiveness, but the presence of the shift at all indicates that these controls should be improved or replaced. It would be necessary to ensure that these shifts could be detected by a control chart monitoring system, and the effect of data aggregation on the ATS performance of control charts

with these properties should be studied. Reynolds and Stoumbos (2004a, 2004b) considered the effects of transient shifts when comparing effective sample sizes.

6.3.3 The Effect of Data Aggregation for Monitoring Process Improvements

Chapters 3 and 4 focus on the detection of decreases in the time-between-events (λ) parameter, or a deterioration of the process. Improvements, or the effectiveness of hazard controls for the safety case, can be noted by the reduction or elimination of control chart signals, but the ATS chart performance for the detection of improvements could also be compared for differing periods of aggregation. Gan (1994) briefly compared the effects for detecting increases in λ , again considering exponential CUSUM charts and Poisson CUSUM charts for 1 and 10 time units of aggregation.

6.3.4 Autocorrelated Data

The assumption of a homogenous Poisson process requires that the count values recorded in each aggregation period are independent of each other. Wu et al. (2010) suggested that multiple injuries occurring from the same catastrophic event can be weighted by the number of affected employees. A similar practice should be considered for future statistical monitoring systems.

6.3.5 Other Distributions

So far, it has been assumed that the time between successive safety incidents will follow an exponential distribution. The validity of this assumption should be tested. A natural alternative to consider would be the Weibull distribution, since the exponential distribution is a specific example of the more general Weibull distribution.

6.4 Value Communication and Usability

In order for any implemented system to be successful, it must be clearly understood by all stakeholders – especially when pairing ideas from two technical fields which don't typically interact, such as occupational safety and statistics. Therefore, it will be necessary for a communication and feedback loop to be developed among the statisticians, practitioners and employees involved in the development and use of such a monitoring system. Various tutorials have been provided for control chart practitioners in other domains, especially healthcare (e.g.,

Mohammed et al., 2008; Benneyan et al., 2003), so it is reasonable to assume that similar guides could be created for safety officials. If the value of monitoring can be effectively communicated, employees will be more likely to participate in the system by reporting incidents, especially leading indicator data which has been historically underreported (Hinze et al., 2013).

An important point of communication between developers and practitioners will be the monitoring system's usability. Before a monitoring system is implemented, its usability should be validated. Thus far, a system safety framework for evaluating the usability of general safety tools has been proposed (Savioja and Norros, 2012), and the usability of statistical monitoring tools has been discussed separately (Reynoso and Olfman, 2004). A combination of these usability evaluations should be considered for use with current and future safety monitoring systems.

References Not Included in Chapters 2-5

- Benneyan, J.C., Lloyd, R.C., and Plsek, P.E. (2003). "Statistical Process Control as a Tool for Research and Healthcare Improvement." *Quality Improvement Research* 12, pp. 458-464.
- Brogmus, G.E. (2007). "Day of the Week Lost Time Occupational Injury Trends in the US by Gender and Industry and their Implications for Work Scheduling." *Ergonomics* 50(3), pp. 446-474.
- Grigg, O. and Farewell, V. (2004). "An Overview of Risk-Adjusted Charts." *Journal of the Royal Statistical Society A* 167(3), pp. 523-539.
- Leamon, T.B. and Murphy, P.L. (1995). "Occupational Slips and Falls: More than a Trivial Problem." *Ergonomics* 38(3), pp. 487-498.
- Lipscomb, H.J., Glazner, J.E., Bondy, J., Guarini, K., and Lezotte, D. (2006). "Injuries from Slips and Trips in Construction." *Applied Ergonomics* 37, pp. 267-274.
- Lovegrove, J., Valencia, O., Treasure, T., Sherlaw-Johnson, C., and Gallivan, S. (1997). "Monitoring the Results of Cardiac Surgery by Variable Life-Adjusted Display." *Lancet* 350, pp. 1128-1130.
- Lovegrove, J., Sherlaw-Johnson, C., Valencia, O., Treasure, T., and Gallivan, S. (1999). "Monitoring the Performance of Cardiac Surgeons." *Journal of the Operational Research Society* 50(7), pp. 684-689.
- Mohammed, M.A., Worthington, P., and Woodall, W.H. (2008). "Plotting Basic Control Charts: Tutorial Notes for Healthcare Practitioners." *Quality and Safety in Healthcare* 17, pp. 137-145.
- Poloniecki, J., Valencia, O., and Littlejohns, P. (1998). "Cumulative Risk Adjusted Mortality Chart for Detecting Changes in Death Rate: Observational Study of Heart Surgery." *British Medical Journal* 316, pp. 1697-1700.
- Reynoso, J.M.G. and Olfman, L. (2004). "An End-User Training Software Tool for Statistical Process Control: An Information Systems Design Theory."
- Savioja, P. and Norros, L. (2012). "Systems Usability Framework for Evaluating Tools in Safety-Critical Work." *Cognition, Technology & Work* DOI: 10.1007/s10111-012-0224-9.
- Sego, L.H., Reynolds, M.R., Jr., and Woodall, W.H. (2009). "Risk-Adjusted Monitoring of Survival Times." *Statistics in Medicine* 28, pp. 1386-1401.
- Steiner, S.H., Cook, R.J., Farewell, V.T., and Treasure, T. (2000). "Monitoring Surgical Performance Using Risk-Adjusted Cumulative Sum Charts." *Biostatistics* 1, pp. 441-452.
- Steiner, S.H., Cook, R.J., and Farewell, V. (2001). "Risk Adjusted Monitoring of Binary Surgical Outcomes." *Medical Decision Making* 21(3), pp. 163-169.

Subramanian, A., Desai, A., Prakash, L., Mital, A., and Mital, A. (2006). "Changing Trends in US Injury Profiles: Revisiting Non-Fatal Occupational Injury Statistics." *Journal of Occupational Rehabilitation* 10(1), pp. 123-155.

Wirtz, A., Nachriner, F., and Rolfes, K. (2011). "Working on Sundays-Effects on Safety, Health, and Work-life Balance." *Chronobiology International* 28(4), pp. 361-370.



Institute of Industrial Engineers 3577 Parkway Lane, Suite 200, Norcross, GA 30092

Content License – Request for Limited Use

Company Virginia Tech, Grado Department of Industrial and Systems Engineering

Address 250 Durham Hall (0118), Blacksburg State VA Zip 24061

Name: Anna Schuh Phone:

Email: schuh@vt.edu

Your project title: "Including Accident Severity in Statistical Monitoring Systems for Occupational Safety"

IIE Content Title:

IIE Publication: 2013 Proceedings of the Industrial and Systems Engineering Research Conference

Month: Year: 2013

The placement of IIE content is for:	Print or electronic 1x use
<input checked="" type="checkbox"/> Professor, Student or IIE Author use 1x placement in paper, course material or instruction	No charge
<input type="checkbox"/> Third party organization using for educational purposes (determined by IIE)	\$25 per page
<input type="checkbox"/> Publisher's 1x use for educational textbook, course book or e-book	\$35 per page
<input type="checkbox"/> Third party organization using for promotional or advertising purposes	\$45 per page

Content needed starts on page: unknown Content needed ends on page: unknown Include Graphics: circle YES or NO

(Not yet in print)

Authorization for payment signature (matches name above)

Anna Schuh

Date: 3/12/2013

By signing this contract, you agree to use the Institute of Industrial Engineers (IIE)'s content for the purposes indicated above and you will not use the content more than once unless you or your company obtains additional rights with a new contract. IIE is the sole the rights granted. Permission is contingent on payment of the content use fee in U.S. dollars. Misuse of this content will be considered breach of this contract and IIE has the right to collect not only 10 xs the fee stated above but any additional damages/payments necessary to compensate for the misuse as well as any legal and court fees incurred by IIE to collect payment for misuse. IIE requires the following verbiage placed on or near where content is used.

"Reprinted with the permission of the Institute of Industrial Engineers, 3577 Parkway Lane, Suite 200, Norcross, GA 30092, www.iienet.org Copyright©2012 by Institute of Industrial Engineers."

Additional Usage Terms: Full payment required before release of content is granted. To ensure that you have content rights please request a receipt for your payment. Receipts of payment will be emailed to the email listed above. Checks or credit cards are accepted in U.S. dollars only.

Thank you for your request.

Please sign and email PDF or fax back to (770) 263-8532.