Development of Safety Event Metrics for an Aviation Organization

Anthony J. Morell, Mary E. Johnson, Edie K. Schmidt, Michael W. Suckow
Purdue University
amorell@purdue.edu, mejohnson@purdue.edu, schmidte@purdue.edu,
msuckow@purdue.edu

Abstract

The introduction of safety management systems (SMS) to the aircraft services industry has stressed the importance of collecting and measuring safety event data to continuously improve management and operational processes. Aircraft maintenance companies provide services to the aviation industry. These companies have begun to collect safety event data and now recognize the need to prepare reports based on this data. While collecting information presents its own difficulties, this study focuses on measuring safety data and presenting information to facilitate management decision making. Aircraft maintenance companies indicate they want a more in-depth analysis of the safety data and anticipate that better decisions may be made by using this data that they have invested in collecting. This study has produced a set of safety event metrics designed to analyze the safety event data and report it in an organized manner to include trends, control charts, Pareto charts, and aging analysis. The results also include aggregated data collected from questionnaires that were administered to managers from the aircraft maintenance industry. These questionnaires were designed to measure each participant's opinion on the usefulness of the set of metrics. Overall, this study has demonstrated that a set of useful metrics can be developed based on the safety event data to support everyday management decisions, as well as provide a foundation for further metric development.

Introduction

Aviation organizations have processes and procedures for collecting data, determining root causes, and recording the findings electronically in their safety event database system. The vast amount of information collected in company safety event databases needs to be distilled into a set of metrics in order for the data to be useful in decision-making. Within the Federal Aviation Administration's (FAA) advisory circular pertaining to safety management systems, it states, "Audits and other information-gathering activities are useful to management only if the information is distilled into a meaningful form and conclusions are drawn to form a bottom line" [1, p.19]. The International Air Transport Association (IATA) states, "To be useful, the data must be transformed into information that can be used by system managers to make informed decisions" [2, p. 4]. Without a useful set of metrics, management and technicians will not be able to identify and implement the proper improvements to processes and procedures.

ISBN 978-1-60643-379-9

This paper discusses the development of a set of useful metrics designed to support management decisions regarding system improvement. These metrics can be useful by providing up-to-date information regarding the safety event data that may be used by management to decide where improvement actions should be focused. The primary objective of the metrics developed was to provide useful information to support decisions made by management and to facilitate improvements in the system. An estimated 80–90 percent of contributing factors are under management control, while 10–20 percent are under the technician's control [3]. By analyzing the data and presenting a useful set of metrics, management has the ability to track and eliminate a large amount of the contributing factors that lead to errors, violations, and subsequent safety events within the workplace. Creating awareness and maximizing learning can be accomplished by sharing findings and recommendations with the affected employees [4]. Management can further reduce the remaining 10–20 percent of contributing factors that are under the technician's control by using these same metrics to increase awareness among the technicians and inspectors.

Review of Literature

Aviation Safety Management Systems

Transport Canada, the International Civil Aviation Organization (ICAO), and the Federal Aviation Administration (FAA) are all globally recognized for their contributions to aviation safety and safety management systems (SMS). In 2005, Transport Canada "placed the conceptual shifts involved in an SMS into the forefront of many airlines' agendas around the world" [5, p. 1]. ICAO's Safety Management Manual is intended to support the implementation of safety management systems [6]. The FAA advises a minimum standard for an aviation SMS [1]. The FAA's safety management standard is parallel to the framework developed by ICAO. The FAA issues and enforces regulations and minimum standards for safety in civil aviation. As of September 2008, the FAA does not require aviation service providers to implement an SMS, but in May 2008, the FAA recommended action to prepare for future implementations of SMS. Updates to FAA regulations may be found at www.faa.gov. Since aviation organizations are extremely safety-conscious, many are beginning to implement their own SMS, prior to a federal regulation requiring one.

A safety management system (SMS) has been defined as "an organized approach to managing safety, including the necessary organizational structures, accountabilities, policies and procedures" [7, p. 1–2]. An SMS is designed to "increase industry accountability, to instill a consistent and positive safety culture, and to help improve the safety performance of air operators. This approach represents a systematic, explicit, and comprehensive process for managing risks to safety" [8]. In three SMS documents [1, 7, 9], data collection and analysis are included as a valuable element of an SMS. Analyzing the data collected from an audit program, investigations, and employee reports allows an organization to be able to evaluate where improvements can be made to the organization's operational processes, as well as the SMS [1].

Investigation of Contributing Factors

The contributing factors to safety events are important to understand. Maintenance Error Decision Aid (MEDA) is an investigative process that was developed to determine contributing factors [10]. "The central philosophy of the MEDA process is that people do not make errors on purpose. While some errors do result from people engaging in behavior they know is risky, errors are often made in situations where the person is actually attempting to do the right thing. In fact, it is possible for others in the same situation to make the same mistake" [11, p. 17].

A large proportion of blame for errors has been traditionally placed on the technician because of the assumption that human error may be attributed to the actions of an individual and not because of the contributions of the environment in which the individual is operating [12]. Reason's Swiss Cheese model showed that events are not caused only by the last event, but that they are actually the end result of a long line of events of which the last act can be hazardous [13]. The MEDA tool was designed to take the investigator beyond the technician's active error and to explore as far up the causal chain as time and money would permit to correct the contributing factors [11].

Safety Metrics

The aviation industry strives for the same goal as any other industry—to reduce the number of events to zero or as close to zero as possible. Measurements of safety performance allow a company to understand system performance and whether or not their safety processes are effective in reducing the amount of events [14]. Measurements may be used to help identify opportunities for improvement [15]. A sound measurement system supports decision making, indicates how the system is performing, helps in establishing priorities on important opportunities for improvement, and verifies that improvements are working [16]. Accident investigation is one area that should be monitored to ensure continuous improvement of the entire safety system [4]. Recommended safety system measurements include the following [4]:

- Percentage where system causes are identified.
- Percentage where causes of human error are identified.
- Percentage of incident reports that are shared with other units.
- Percentage of follow-up actions and learning shared.
- Percentage of incidents investigated to root causes.
- Average time from incident to investigation completed.
- Average time from incident completion to correction.
- Percentage of investigations that show planning failure
- Percentage of accident reviews with leadership participation.

Allocating scarce resources is a challenge management faces on a daily basis. The Pareto chart is one way to make educated decisions based on data analysis. The Pareto concept, also known as the 80–20 rule, was developed by Vincent Pareto and has demonstrated that 20 percent of the known variables will account for 80 percent of the results [17]. Pareto charts may be used to identify the large problems that may be reduced more quickly and with greater impact, as opposed to focusing on eliminating small problems [18]. Using a Pareto chart, management can see the arrangement of data (errors, defects, or failures), view the most frequent deficiencies of a system, and eliminate or reduce these items as much as possible.

Some researchers have adapted quality management practices to fit safety management's needs. In one such study, Pareto charts were used to indicate the frequency, severity, and location of problems in a facility [19]. This data was combined with perception survey data (proactive measures) to obtain a more complete view of what was occurring in the workplace [19].

Developing measurement systems is difficult, complex, and important [15]. When developing the set of metrics, the flow of the information through the system should also be addressed. Transport Canada [20] has documented the type of information that should be stored in such a database, as well as its path through the improvement process. The process map, shown in Figure 1, shows how both reactive and proactive data flow through the same path into a database. Data should then be analyzed and the results communicated throughout the facility as part of a continuous system evaluation. System evaluation should be a continuous loop of information facilitating improvement to the SMS. In addition, measurements should be expected to change over time as conditions change [15]. By examining the literature, the authors defined the characteristics of useful metrics (as shown in Table 1).

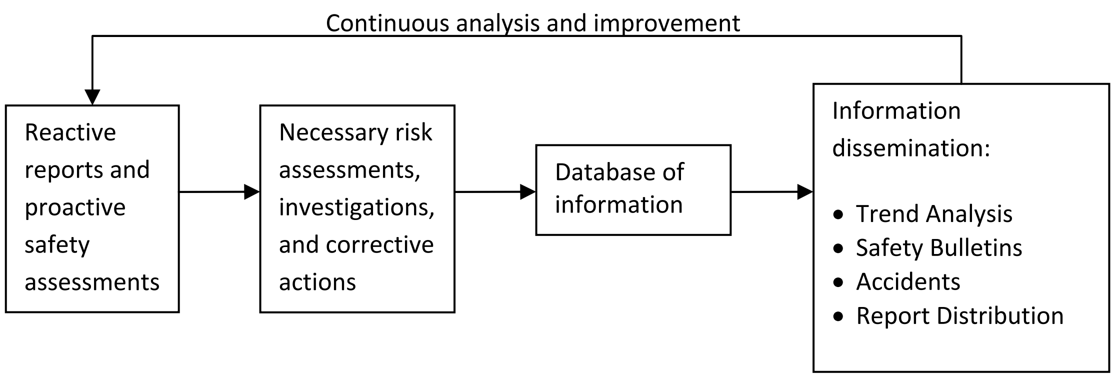

Figure 1: SMS Process Flow (adapted from [20])

Table 1. Characteristics of Useful Metrics

Characteristics of Useful Metrics
Easy to understand [10]
Supports decision making [2, 15]
Captures opportunities for improvement [1, 4, 15, 16]
Provides understanding of system performance [1, 14, 16]
Isolates important issues [16, 18]

Methodology

The goal of this project was to design and test a set of useful metrics based on analysis of the types of safety event data stored in a typical aviation SMS. These metrics are designed to facilitate management decisions in the continual improvement of safety and operational procedures at an aviation organization. A four-step process was followed in the development of these metrics.

<u>Step 1: Understand What Data Is Available</u>

The data available in an SMS database largely determines what data can be presented or analyzed. Typical databases collect information on the date, type of incident, location of incident, and other data recommended by standard approaches, such as MEDA.

<u>Step 2: Develop Metric Usefulness Questionnaire</u>

A questionnaire was developed to evaluate each metric for usefulness. The questions were based on the characteristics in Figure 1. The questionnaire consisted of questions related to the characteristics of useful metrics, with scoring based on a five-point Likert scale.

<u>Step 3: Develop Metrics</u>

A set of metrics was developed using the data available in SMS databases. The metrics were selected based on those that had the potential to convey trends and provide other potentially useful information.

<u>Step 4: Validate Usefulness of the Metrics</u>

The metrics developed in Step 3 were presented to eight management-level aircraft maintenance industry experts for evaluation. During the review of the metrics, these managers participated in a discussion and question/answer period. Finally, the experts were given an opportunity to voluntarily answer the questionnaire developed in Step 2 to evaluate the usefulness of the metrics. Participants, who were all from the same corporation, were instructed not to identify themselves on the questionnaires. The corporation had an SMS database but no set of metrics developed using the information available in the database.

Metric Development

Through multiple discussions with industry experts from the areas of quality, safety, and human factors, a total of eight metrics were chosen to support the decisions of managers and to drive system improvement efforts into the correct areas. The eight metrics are:

1. Aging chart
2. Events opened/month
3. Events closed/month
4. Response time
5. Corrective action time

6. Frequency of each event type
7. Frequency of severity levels/month
8. Frequency of causal factors/month

The developed metrics were limited by the categories contained in the safety event database system. This section describes each of these metrics, what charts were used, and how the set of metrics can be used together to provide an illustration of the safety event data. No single metric is presented as describing the overall performance of the organization. As a group, these metrics provide insight into organizational performance using SMS data.

Aging Chart

An aging chart was selected to display the number of days that events have been open for all open events in the system. The chart is formatted as a histogram with boundaries set for 5-day intervals, as seen in Figure 2.

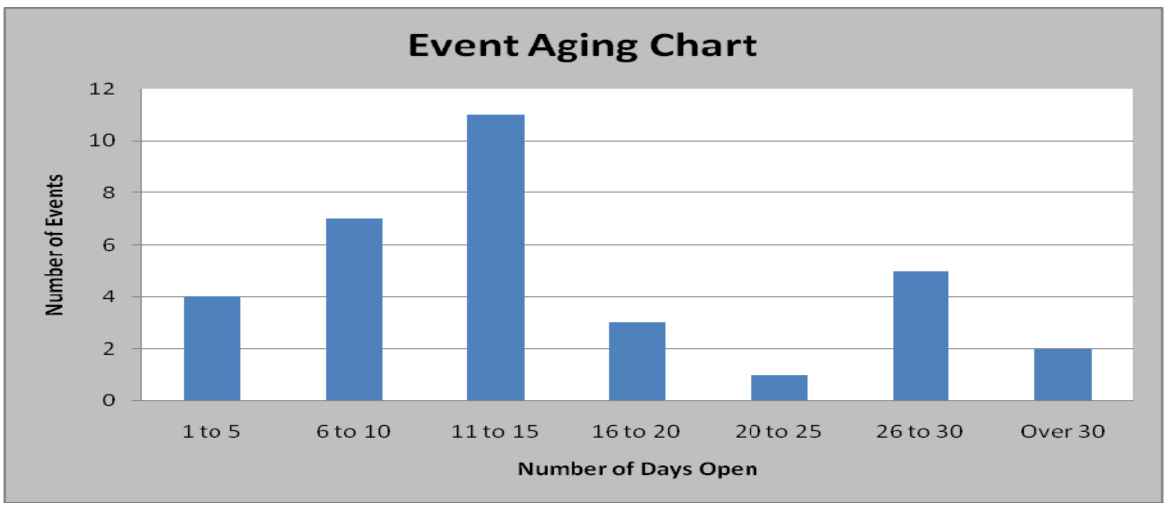

Figure 2. Event Aging Chart

This type of information could become a valuable tool in identifying those events that have exceeded a time limit set by management. To develop a limit, management will need to allocate appropriate resources to quickly resolve events and monitor the aging charts for a given period of time to determine where the standard should be set.

Events Opened/Closed per Month

A run chart and control chart were selected to provide information to managers regarding organization performance. Separate run charts for the number of events opened each month and the number of events closed each month were developed. These charts may be analyzed to provide evidence of any patterns within the discrete data, such as trends, oscillations, mixtures, and clustering. Examples of discrete data may include the number of complaints, the number of defects, or, in this case, the number of events opened per month, as shown in

Figure 3. The run chart may be set to display the number of events over specified time periods such as years, months, quarters, or weeks. This allows management to tailor the chart to the time period of interest. As more data is collected, the sophistication of analysis may also include statistical process control to identify common cause and special cause variation.

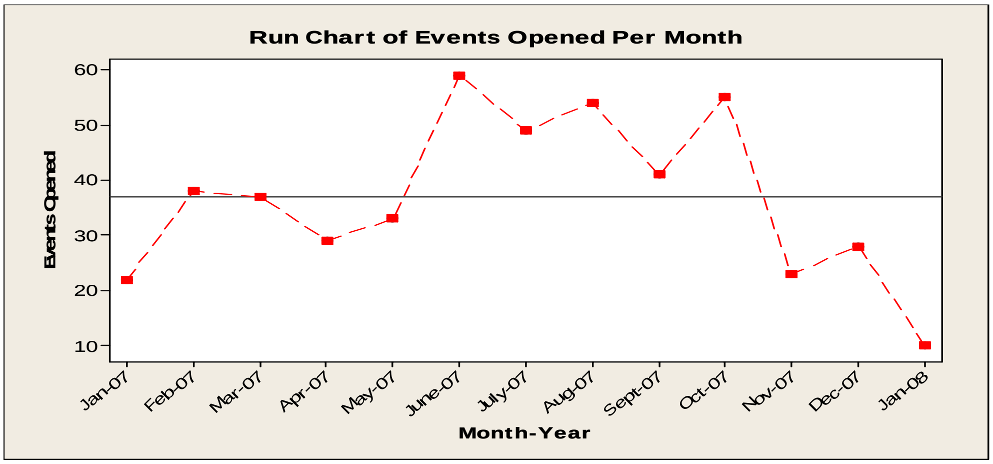

Figure 3. Run Chart

<u>Response Time</u>

Frequency histograms were selected for response time data presentation, as shown in Figure 4. Response time can be defined as the number of days between the “Response due date” and the “Date response entered.” A negative number represents a response that was entered before the due date, while a positive number means the response was late. The goal of this measurement is to provide managers with an average on how accurately an event response due date is being predicted. Furthermore, the information in these charts could provide insight into why events are not being responded to on time or what is causing the responses to be early. The response data is charted in the same manner as the corrective action time metric, using a histogram.

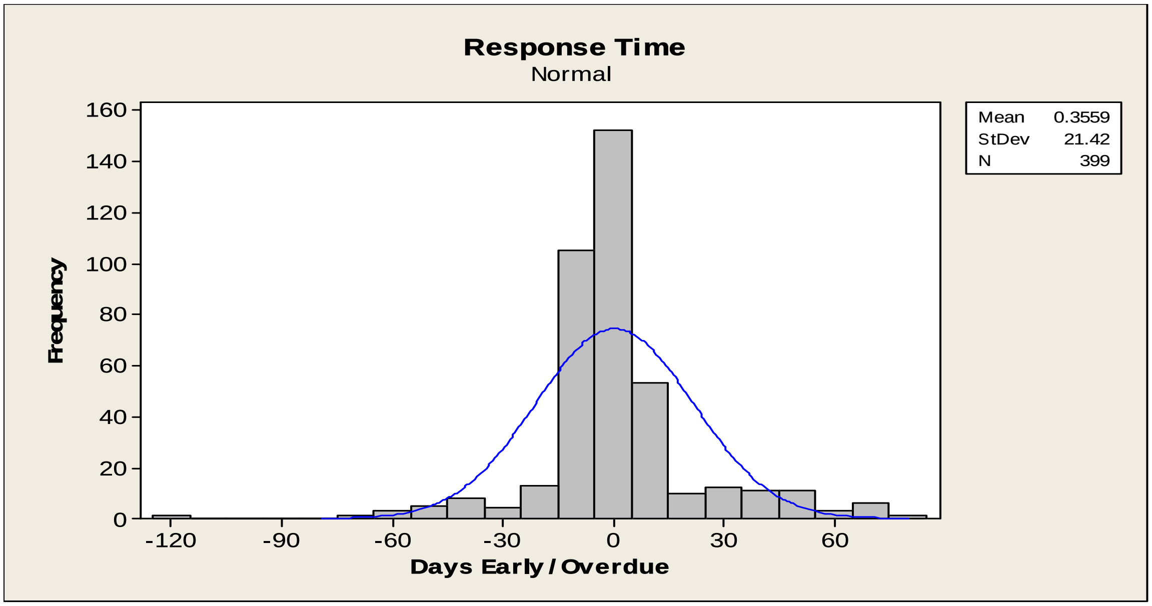

Figure 4. Response Time Histogram Chart

Corrective Action Time

Frequency histograms were selected for corrective action time charts. These charts are similar in structure to those shown in Figure 4. Corrective action time is defined in this study as the number of days between the "Event Entered" date and the "Event Closed" date. Since an event can only be corrected in a positive amount of time, the charts produced will only indicate a positive number to the nearest day. In addition to displaying the distribution, the data may also be used to calculate other statistical measures such as average, median, and standard deviation.

Frequency of Each Event Type

A Pareto chart was selected to display the frequency of each event type for a given month. This information is displayed by month in Figure 5. One could use the Pareto chart to identify the most frequent event type, and then view the run charts or control charts for the same time period to better understand system performance. The event types can be charted over the entire time span of the database or by a given time period, such as months, years, or quarters, depending on the needs of management. Following the 80–20 rule, these charts identify the important few events to focus improvement efforts.

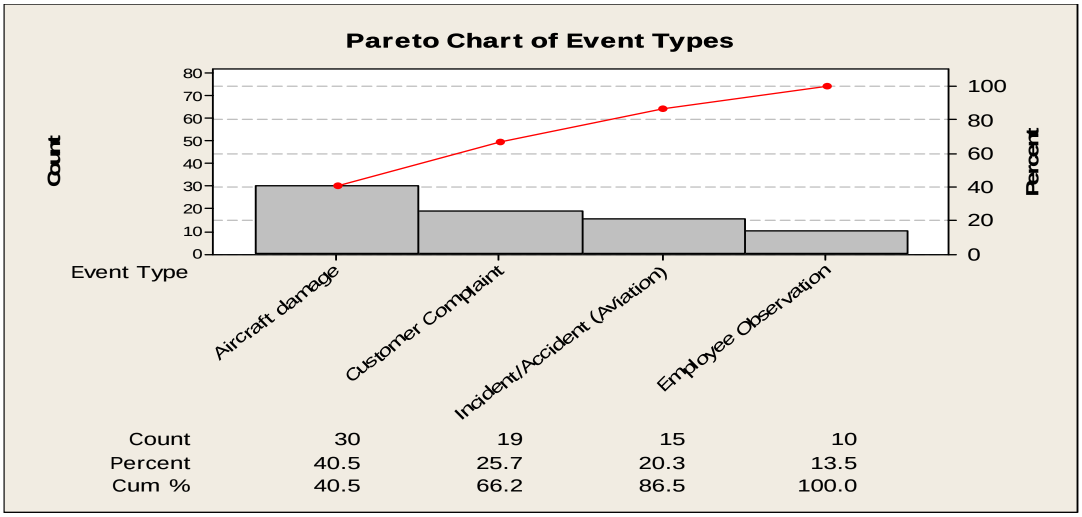

Figure 5. Pareto Chart

The preceding six charts provide a general picture of the system. The remaining charts are meant to be used to gather additional information to clarify any questions of the data.

Frequency of Event Severity Levels per Month

A Pareto chart was selected to provide more detailed information regarding the type of severity levels seen within the database over a certain time period (in this case per month). In Figure 6, the most severe level is “D,” while the least severe is “A.” Tying this chart to the Pareto for event types allows management to view the distribution of severity levels in relation to the events occurring during the same time period. Management should be careful not to ignore the most severe events just because they may be the least frequent. Events that are of severe levels should always be investigated.

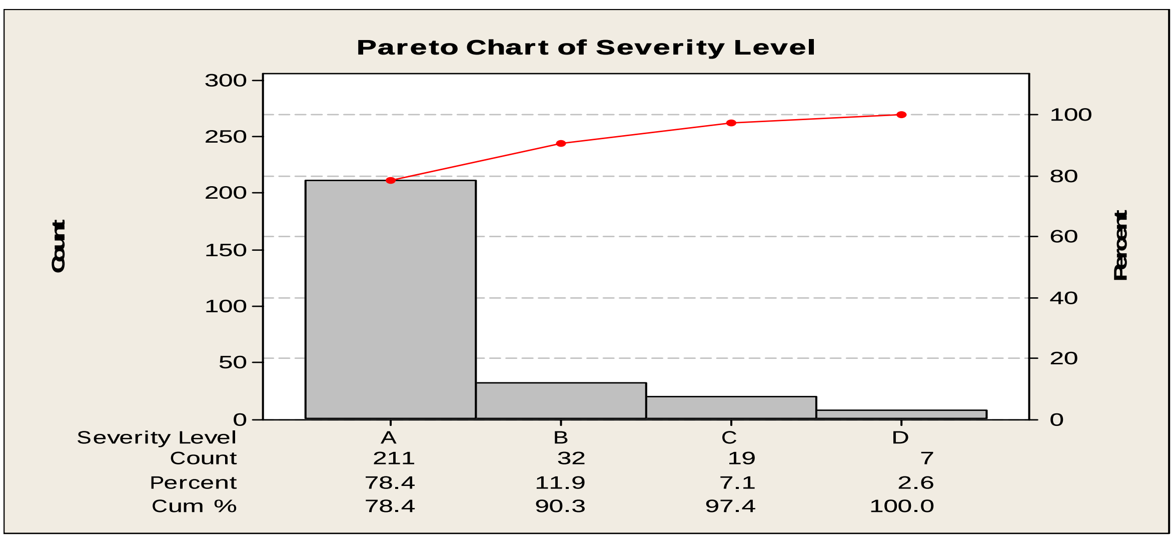

Figure 6. Pareto Chart of Severity Level

Frequency of Each Level of Causal Factors per Month

A Pareto chart was selected to display the frequencies of the first two causal factor categories, similar to the chart in Figure 6. Investigations into root cause should identify the causal factors for events. Each causal factor is comprised of three levels of categories: factor class, causal factor type, and response. The charts can be organized in a multitude of ways, depending on what is being investigated. In this study, the Pareto charts display the overall frequencies at each category level. Even at this level of granularity, management may begin to use these charts to understand which root causes are most common. This allows management to focus improvement efforts on the areas that may lead to greater improvements in a shorter period of time.

Combining the Metrics

When viewed independently, these charts provide information regarding measured aspects of the safety and quality systems. Though valuable as independent sources of information, the advantage of these metrics is that they may be used together to support management decisions. An out of control data point on statistical process control charts for events opened or closed per month can drive a manager to investigate the type of events, their severity levels, and what type of causal factors were found. Furthermore, the corrective action time and response accuracy can be tied in with the severity level and event type to determine why corrective action times and responses were at certain levels. The information discovered during these investigations can be used to improve processes, provide foundations for new standards, create new measurements, and aid in the development of new processes in the future.

Usefulness of Metrics

Once the analysis of the safety event data and the development of the metrics concluded, the metrics and sample charts were presented to a group of eight aviation industry experts. The attendees represented management from areas including quality, safety, human factors, and information technology. During the presentation, the managers began to discuss ideas such as how to more accurately identify the categories of causal factors and event types in their data collection.

A total of eight managers participated in the questionnaire. The results for each question were collected and are reported in aggregate in Table 2. The overall mean score was 4.38, with a median of 4.5 on a scale of 1 to 5, with 5 being “strongly agree.” These scores indicate that the respondents agreed that the metrics produced in this study contained useful information. Question 5 had the highest mean score at 4.63, while question 6 scored the lowest at 4. A score less than four was assigned by four respondents, three times for question 6 and once for question 2. Since the metrics produced in this study were the first set of metrics created for the SMS and seen by these managers, it is understandable if these managers may not immediately declare to use them on a regular basis.

Table 2. Usefulness Questionnaire and Aggregated Results

Question	Mean	Median
1. The metrics were easy to understand.	4.50	4.5
2. I am able to make decisions supported by these metrics.	4.25	4
3. The metrics identify opportunities for improvement.	4.50	4.5
4. The metrics assist in displaying the level of system performance.	4.38	4.5
5. The metrics identify important issues.	4.63	5
6. I would use these metrics on a regular basis.	4.00	4
7. These metrics should be included in reports.	4.38	4

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5

Conclusions

This study sought to develop a set of useful SMS metrics through the review of previous literature in the areas of data collection, data analysis, and existing safety measures, as well as discussions with aviation industry experts. The outcome is a set of useful metrics that have already begun to support management decisions at one company who chose to adopt them. This study did not seek to develop a comprehensive set of metrics. This study has provided a foundation for future research and development of safety event data measurements.

The primary objective of the metrics produced from this study was to provide useful information to various levels of management. These metrics are meant to support decisions made by management and to facilitate improvements in the system. During the presentation of the metrics to the group of industry experts, the usefulness of the metrics became apparent. During the presentation, the managers began discussing ideas on how to incorporate these metrics into more localized measures, such as scorecards for each work area. Based on the metrics presented, the managers also discussed how to gather more detailed categories of causal factors and event types in their data collection.

The usefulness of the set of metrics developed in this study is supported by the results of a voluntary questionnaire. The experts who participated in the metric presentation were given an opportunity to answer the questionnaire and provide their opinion on the usefulness of the developed metrics. Respondents represented management from areas including quality, safety, human factors, and information technology. The results of the questionnaires confirm that a set of useful metrics can be developed from the analysis of safety event data to support management decisions.

The literature reviewed described the common elements of various safety management systems, how root causes are determined, and the characteristics of useful metrics. Based on the analysis of the safety event data and the results of the questionnaire, a safety event database can produce useful metrics to support management decisions.

Recommendations

This paper assumes that organizations have taken the first step in developing useful metrics by collecting data from every safety event, regardless of the severity of the event. The developed metrics were determined to be useful according to a group of aviation experts who rely on information to make decisions. As aviation organizations begin to understand what type of data is useful and make necessary changes to their database systems, the organizations should continue to develop metrics to help management make decisions.

By identifying the specific work area or product type, the data could be used to provide more specific information regarding safety events. This information could allow managers to pinpoint the events by focusing on a specific area of the facility. In addition to localizing events, placing specific weights on certain types of information could provide a more useful metric. One example would be to weigh the events per time period by the number of labor-hours expended. Labor-hours are a common measurement and are already collected by most aviation organizations. Many times the labor-hour information is stored on another database and would require additional programming to calculate the number of events per man-hour. More detailed charts could be created for the aging charts, with specific aging charts for each event type or severity level. This could help management gain an understanding of the trends for each category and allow them to create a baseline for the amount of time an event should remain open.

As the metric systems are developed and improved for safety management systems, the addition of analyzing severity and risk of each event type could prove to be invaluable. Severity and risk are already a category in many SMS databases. Any new metric should go through an approval process similar to this study's questionnaire. A questionnaire process should involve representatives of any group that will rely on the information if the metric is implemented. Creating a metric approval process to guarantee its usefulness will lead to a useful set of measurements, with each metric serving a purpose to the managers who will rely on the information to make decisions.

References

[1] Federal Aviation Administration, "Introduction to safety management systems for air operators," *FAA Advisory Circular*, No. 120–92, 2006, Retrieved October 14, 2007,

http://rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/0/6485143d5ec81aae8625719b0055c9e5/$FILE/AC%20120-92.pdf.

[2] IATA, “Global Aviation Safety Roadmap,” 2007, Retrieved November 4, 2007, http://www.iata.org/NR/rdonlyres/D440B2E7-1DE8-4130-A796-F66BB5DDB451/0/GlobalAviationSafetyRoadmap.pdf.

[3] Boeing, “Maintenance Error Decision Aid,” 2005, Retrieved October 1, 2007, http://www2.hf.faa.gov/opsManual/assets/pdfs/MEDA_guide.pdf.

[4] Petersen, D., *Measurement of Safety Performance*, Des Plaines, Ill: American Society of Safety Engineers, 2005.

[5] FSF Editorial Staff, “Unlocking the potential of a safety management system,” *Flight Safety Digest*, Vol. 24, No. 11–12, 2005, pp. 1–15.

[6] ICAO, “Welcome to the safety management website,” 2007, Retrieved December 5, 2007, http://www.icao.int/anb/safetymanagement/.

[7] ICAO, *Safety Management Manual*, 2006, Retrieved October 1, 2007, http://www.icao.int/icaonet/dcs/9859/9859_1ed_en.pdf.

[8] Transport Canada, “Regulatory amendments to improve aviation safety finalized,” 2005, Retrieved December 3, 2007, http://www.tc.gc.ca/mediaroom/releases/nat/2005/05-h142e.htm.

[9] Transport Canada, “Components of the safety management system,” CAR Part VII, Section 705.152, 2005, Retrieved December 4, 2007, http://www.tc.gc.ca/civilaviation/regserv/affairs/cars/part7/705.htm.

[10] Rankin, W., Hibit, R., Allen, J., and Sargent, R., “Development and evaluation of the maintenance error decision aid (MEDA) process,” *International Journal of Industrial Ergonomics*, 26, 2000, pp. 261–276.

[11] Rankin, W., “MEDA Investigation Process,” *Boeing Commercial Aero*, 2nd Quarter, 2007, Retrieved September 18, 2007, http://www.boeing.com/commercial/aeromagazine/articles/qtr_2_07/AERO_Q207_article3.pdf.

[12] Reason, J., *Human Error*, Cambridge University Press: New York, 1990.

[13] Reason, J., *Managing the Risks of Organizational Accidents*, Ashgate: Vermont, 1997.

[14] Toellner, J. “Improving safety and health performance: Identifying and measuring leading indicators,” *Professional Safety*, Vol. 46, No. 9, 2001, pp. 42–47.

[15] Sink, D.S., “The role of measurement in achieving world class quality and productivity management,” *Industrial Engineering*, Vol. 23, No. 6, 1991, pp. 23–28, 70.

[16] Sink, D.S., and Smith, G.L., “Reclaiming process measurement,” *IIE Solutions*, Vol. 31, No. 2, 1999, pp. 41–46.

[17] Basile, F., “Great management ideas can work for you,” *Indianapolis Business Journal*, Vol. 16, 1996, pp. 53–54.

[18] Eckes, G., *The Six Sigma Revolution: How General Electric and Others Turned Process into Profits*, New York: John Wiley, 2001, p. 85.

[19] Williamsen, M.M., “Six Sigma safety: Applying quality management principles to foster a zero-injury safety culture,” *Professional Safety*, Vol. 50, No. 6, 2005, pp. 41–49.

[20] Transport Canada, “Safety management systems for flight operations and aircraft maintenance organizations (TP 13881E),” 2002, Retrieved October 13, 2007, http://www.tc.gc.ca/publications/EN/TP13881/PDF%5CHR/TP13881E.PDF.

Biography

ANTHONY J. MORELL was a graduate student in the Aviation Technology department at Purdue University. His BS and MS degrees are from Purdue University.

MARY E. JOHNSON is an Associate Professor in the Aviation Technology department and the Industrial Technology department at Purdue University. She is interested in performance measurement and improvement and the incorporation of creativity in the learning process. Her PhD is in Industrial Engineering from The University of Texas at Arlington.

EDIE K. SCHMIDT is an Associate Professor in the Industrial Technology department at Purdue University, where she conducts research in supply chain, distribution, and project management. Her PhD is from Purdue University.

MICHAEL W. SUCKOW is an Assistant Professor in the Aviation Technology department at Purdue University. He is a pilot and has extensive airline experience prior to joining Purdue.