ENHANCING SAFETY PERFORMANCE WITH A LEADING INDICATORS PROGRAM

C M Tomlinson, ABS, UK
B N Craig, Lamar University, USA
M J Meehan, AP Moller-Maersk, Denmark

SUMMARY

Safety performance monitoring through leading indicators is a key initiative that may be able to improve safety performance. Leading indicators are safety metrics that are associated with, and precede, an undesirable/unexpected consequence such as an operational incident, near miss or personal injury. Their utility for risk management comes from the possibility that they may reveal areas of weakness in advance of adverse events.

This paper presents the results of research undertaken by ABS and Lamar University with support from AP Moller-Maersk. It summarizes the development of the safety culture and leading indicators initiative by ABS and details a method whereby marine organizations with cargo-carrying commercial vessels can develop their own leading indicator programs.

Two approaches to identifying leading indicators are presented: First, from safety metrics data and second, using the results from a safety culture survey. The paper discusses the use of metrics, safety performance data, safety factors and data analysis, and provides guidance on how to incorporate the results into an organization’s continual improvement program.

1. INTRODUCTION

Safety performance has traditionally been monitored by ‘after the loss’ measures to assess outcomes such as accident and injury rates, incidents, and dollar costs. These are known as lagging indicators. For the last two decades there has been a growing recognition across various sectors that data from lagging indicators is limited. It comes too late to allow for preventative action to be taken, and all too often offers little insight into how to prevent further losses.

1.1 LAGGING INDICATORS OF SAFETY

Lagging indicators give a snapshot, or update, of performance but do not give any indication of future results, or if the present results are sustainable [1]. Lagging indicators characteristically:

- identify trends in past performance
- assess outcomes and occurrences
- have a long history of use, and so are an accepted standard
- are relatively easy to identify and analyze

In the aftermath of catastrophes, it is common to find prior indicators, missed signals, and dismissed alerts which, if they had been appropriately addressed at the time of identification may have averted the disaster. Lagging indicators fail to draw attention to these alerts and signals.

Ideally, what is required is a set of leading indicators that can predict future performance so that interventions can be made before accidents or incidents occur [2].

1.2 LEADING INDICATORS OF SAFETY

Over the past two decades, improved safety performance has been associated with a number of measurable activities in various industries, opening up the possibility that some of these metrics may be leading indicators for safety performance. The National Academy of Engineering defines leading indicators as conditions, events, and sequences that precede and lead up to accidents [3]. They must also have some value in predicting the arrival of the event, whether it is an accident, incident, near miss, or undesirable safety state [4].

Examples of leading indicator programs developed in non-marine sectors include: hazard identification and analysis for offshore oil and gas [5]; indicators for the energy and related process industries [6]; accident precursor assessment programs in nuclear safety [7, 8]. Leading indicators can:

- reveal areas of weakness in advance of adverse events
- be associated with proactive activities that identify hazards
- aid risk assessment and management
- complement the use of lagging indicators by compensating for their shortcomings [5]

For leading indicators to play an effective role in the improvement process, there must be an association between the inputs that the leading indicators are measuring and the desired lagging outputs [5], and leading indicators should indicate the direction of future lagging results [1]. Examples of metrics that could be leading indicators are: the size of the safety budget,
safety audit scores, the number of safety inspections, and the number of safety meetings involving management. Leading indicators are leading (as opposed to lagging) measures, and leading in the sense that they are the prime metrics associated with safety performance for a particular organization.

1.3 KEY PERFORMANCE INDICATORS

Leading indicators are frequently confused with key performance indicators (KPIs). KPIs are associated with organizational performance which may, or may not, be safety-related. Examples of KPIs are: budgetary control per vessel; dry-docking planning performance, and vessel availability [9]. KPIs may be leading or lagging indicators. In contrast, leading indicators of safety are always associated with safety performance.

2. DEVELOPMENT OF THE INITIATIVE

For some time, ABS has been investigating a method for identifying potential leading indicators of safety. Beginning in 2003, initial feasibility research was conducted at Rensselaer Polytechnic Institute USA, with assistance from Virginia Commonwealth University. This stage of the research established the viability of identifying statistical correlations between leading indicators and safety performance data.

The research undertaken in the initial phase was used as the basis for the initiative developed at ABS and Lamar University. During the development phase, four case studies were undertaken with marine organizations:

- a domestic U.S. tanker organization
- an international tanker organization
- a domestic U.S. container and government shipping organization
- a large international container and tanker organization (AP Moller-Maersk)

2.1 THE AP MOLLER-MAERSK STUDY

This study began in July 2008 with two objectives:

- to identify and analyze the container fleet’s leading indicators of safety
- to investigate the quality of APMM’s safety culture

Subjective safety culture data was gathered from forty shore side personnel in offices in Copenhagen, Singapore, Cape Town and Rotterdam, and from approximately eight hundred shipboard personnel onboard one hundred and ten ships. The safety culture questionnaire contained items on shipboard and shore side operations, occupational safety and health, and individuals’ jobs. Demographic data was also collected such as nationality, age, experience in current position, experience with the company, and experience in marine industry. Statistical data analysis was performed and differences in safety culture were identified based on age, gender, job title, nationality, and experience.

In early 2009, safety metrics and safety performance data were accessed from company records for the previous six years in order to perform the leading indicators of safety analysis. This was done by correlating the company’s safety metrics with its safety performance data over the preceding years. Safety performance data included personnel health and safety data as well as operational incidents. Note that negative correlations were expected. For example, as the number of safety inspections increased, the number of operational incidents was expected to decrease. The following leading indicators of safety analyses were assessed:

- organizational metrics vs. organizational safety performance for the same year
- organizational metrics vs. one-year delayed organizational safety performance
- organizational metrics vs. two-years delayed organizational safety performance
- shipboard questionnaire vs. shipboard safety performance

2.2 AP MOLLER-MAERSK STUDY RESULTS

An analysis of organizational safety metrics and safety performance data revealed that a subset of these metrics had a significant association (strong negative correlation) with safety performance.

2.2(a) Same-year analysis

For the same-year analyses of metrics and safety performance data, the significant associations were:

- number of safety management meetings (2003 – 2008) vs. restricted work accident frequency (2003 – 2008) \[r = -0.886, p = 0.019\]
- percentage of incident reports on which root cause analysis was undertaken (2003 – 2008) vs. restricted work accident frequency (2003 – 2008) \[r = -0.943, p = 0.005\]
- number of safety inspections vs. restricted work accident frequency (2003 – 2008) \[r = -0.886, p = 0.019\]
- percentage of incident reports on which root cause analysis was undertaken (2003 – 2008) vs. total recordable frequency (2003 – 2008) \[r = -0.886, p = 0.019\]
- percentage of incident reports on which root cause analysis was undertaken (2004 – 2008) vs. restricted work accident frequency (2004 – 2008) \[r = -0.900, p = 0.037\]
- percentage of incident reports on which root cause analysis was undertaken (2004 – 2008) vs. total recordable frequency (2004 – 2008) \[r = -0.900, p = 0.037\]
The analyses resulted in identical r-values because Restricted Work Accident Frequency is a subset of Total Recordable Frequency and there is a small sample size (five-six years).

An example of the strong negative correlation for the same year analysis is shown in Figure 1. The Y-axis on the left of the graph indicates the percentage of incident reports on which root cause analysis was undertaken for 2003 through 2008, and the right Y-axis indicates the restricted work accident frequency from 2003 to 2008.

The example graph in Figure 1 shows the increasing percentage of incident reports resulting in a root cause analysis (from 22% to 47%) was associated with a decreasing restricted work injury case frequency (from 4.7 to 1.8) in the years 2003 to 2008. Similar negative associations were found for the other bulleted items.

![Figure 1: Percentage of Incident Reports on Which Root Cause Analysis was Undertaken (2003 – 2008) vs. Restricted Work Accident Frequency (2003 – 2008) – Same Year](image1.png)

Figure 2: Number of Safety Performance Indicators Utilized (2003 – 2007) vs. Total Recordable Accident Frequency (2004 – 2008) – One Preceding Year

2.2(b) One-year delayed analysis

Analysis was also undertaken on the relationship between safety metrics of one year with safety performance in the following year. Significant results were found for:

- number of safety performance indicators utilized (2003 – 2007) vs. restricted work accident frequency (2004 – 2008) \( r = -0.949, p = 0.014 \)
- number of safety performance indicators utilized (2003 – 2007) vs. total recordable frequency (2004 – 2008) \( r = -0.949, p = 0.014 \)

The analyses resulted in identical r-values because Restricted Work Accident Frequency is a subset of Total Recordable Frequency and there was a small sample size (five years).

An example of this strong negative correlation in the one preceding year analysis is shown in Figure 2. The Y-axis on the left of the graph indicates the number of safety performance indicators utilized for 2003 through 2007, and the right Y-axis indicates the total recordable accident frequency from 2004 to 2008.

The example graph in Figure 2 shows the increasing number of safety performance indicators utilized for the years 2003 to 2007 (from 4 to 7) was associated with a decreasing total recordable injury case frequency for the years 2004 to 2008 (from 5.7 to 3.5).

![Figure 2: Number of Safety Performance Indicators Utilized (2003 – 2007) vs. Total Recordable Accident Frequency (2004 – 2008) – One Preceding Year](image2.png)

2.2(c) Two-years delayed analysis

Analysis was also undertaken on the relationship between safety metrics of one year with safety performance two years later. Significant results were found for:

- percentage of incident reports on which root cause analysis was undertaken (2003 – 2006) vs. restricted work accident frequency (2005 – 2008) \( r = -1.000, p < 0.001 \)
- percentage of incident reports on which root cause analysis was undertaken (2003 – 2006) vs. total recordable frequency (2005 – 2008) \( r = -1.000, p < 0.01 \)

Again, the analyses resulted in identical r-values because Restricted Work Accident Frequency is a subset of Total Recordable Frequency and there was a small sample size (four years).

An example of this strong negative correlation in the two years delayed analysis is shown in Figure 3. The Y-axis on the left of the graph indicates the percentage of incident reports on which root cause analysis was

Figure 3 shows the increasing percentage of incident reports on which root cause analysis was undertaken for the years 2003 to 2006 (from 22 to 37) was associated with a decreasing total recordable injury case frequency for the years 2005 to 2008 (from 4.6 to 1.8).

**Figure 3:** Percentage of Incident Reports on Which Root Cause Analysis was Undertaken (2003 – 2006) vs. Restricted Work Accident Frequency (2005 – 2008) – Two Preceding Years

These results served to validate the research approach taken. Several lessons were learnt from the case study, including:

- the desirability of developing a metrics hierarchy - when it became apparent that not all metrics are equally useful for a leading indicators exercise for all organizations (see section 4.3 for full details)
- the expansion of the method to cover metrics kept at the vessel level and not held centrally
- the need for computerised support for organizations wishing to self-assess their leading indicators – the statistical analysis is not particularly difficult, but it is onerous
- the research effort should provide detailed guidance on how to use the results

AP Moller-Maersk gained sufficient confidence in the approach taken, and the results obtained, that it has continued to collaborate with the development of the ABS leading indicators initiative by providing user requirements for the computerised assistance now being developed (see section 8). Full details of the AP Moller-Maersk safety culture results (shipboard vs. shore side) have been published elsewhere [10].

3. **THE ABS MODEL**

The model shown in Figure 4 indicates that there are several approaches to trying to improve safety performance by improving social and organizational aspects of the company.

**Figure 4:** ABS Safety Culture and Leading Indicators Model

The most basic, but time-consuming, approach is to conduct a safety culture assessment and to act on the results. This could be done as a stand-alone assessment or it could be carried out in conjunction with a leading indicators process.

There are two ways for conducting the leading indicators process:

- Identifying *objective* leading indicators. This is done by correlating safety metrics with safety performance data. This is the preferred approach because of its objectivity; because it utilizes metrics that the organization has collected; and because it does not require a survey of the workforce, which can be time-consuming. This can be done at three levels:  
  - at the organizational level  
  - across business units  
  - across the fleet
- Identifying *subjective* leading indicators from the results of a safety culture survey. These indicators are based on the values, attitudes, and observations of employees. This method may identify beneficial safety metrics not yet tracked by the organization. This approach may be used...
when the organization lacks sufficient metrics to use the objective leading indicators process.

Note that there are a number of criteria for undertaking a leading indicators program and for each type of assessment. For example, to undertake the organizational level analysis, the organization must have been collecting safety metrics for at least five years.

Although the ABS model is generic it has only been applied to marine organizations with cargo-carrying vessels. Some aspects of the toolkit, such as the safety culture questionnaires, would require tailoring for other types of commercial vessels.

4. **A LEADING INDICATORS PROGRAM**

The purpose of a leading indicators program is to identify which safety metrics are strongly associated with safety performance in a particular organization. This information can be used to guide actions to improve future safety performance. This section introduces the basic concepts and principles of a leading indicators program that organizations can use to self-assess their potential leading indicators of safety.

4.1 **GENERAL CRITERIA FOR UNDERTAKING A LEADING INDICATORS PROGRAM**

The leading indicators approach to improving safety performance is likely to be more effective when the technical aspects of safety are performing adequately and the majority of operational incidents and accidents appear to be due to human error or organizational factors. Organizations should be considering a leading indicators approach if the following criteria are met:

- the organization is compliant with all relevant regulations
- the organization has a genuine desire to prevent operational incidents and personal injuries and is not solely driven by statutory compliance
- the organization is relatively stable, not in the middle of mergers, acquisitions or significant reorganizations

If an organization does not meet these criteria, then it may not be ready for a leading indicators program.

4.2 **ASSESSMENT CRITERIA**

In addition, the organization should also meet one of the following criteria, depending on which leading indicators assessment is to be undertaken:

- an objective leading indicators assessment requires that safety metrics have been collected for a period of time, at least five years for an organizational level analysis, and at least one year for the business unit or fleet level
- a subjective leading indicators assessment requires that a safety culture survey is performed and the results utilized

4.3 **SAFETY METRICS**

Objective leading indicators are identified by correlating safety metrics with safety performance data. ABS research has identified three types of metrics that have different levels of usefulness for inclusion in a leading indicators program, shown in Figure 5.

![Figure 5: The Metrics Hierarchy](image)

4.3(a) **Baseline Metrics**

Baseline metrics form the foundation of a safety culture and should be collected. However, because they are expressed as the presence or absence of an activity, procedure or policy (and not as interval data, ratios, frequencies, etc. that can vary) they are unsuitable for inclusion in a leading indicators program. Examples of baseline metrics are:

- provision of a communications training program
- presence of a crew feedback system concerning near misses and hazard identifications
- establishment of a fair system for incident investigation
- presence of a maintenance budget

4.3(b) **Subsidiary Metrics**

Subsidiary metrics are useful in a leading indicators program until they peak or become invariant, which they may do as the safety culture takes root. For example, once “Percentage of crew who have PPE” consistently attains 100%, it is no longer useful as a metric for correlating with safety performance. Examples of subsidiary metrics are:

- percentage of employees receiving ALL safety training
- number of safety inspections per annum
- frequency of safety meetings attended by senior management
• number of safety performance indicators utilized.

4.3(c) Core Metrics

The core set of metrics are eminently suitable for inclusion in a leading indicators program by all organizations, even those with a mature safety culture. Examples of core metrics are:

- percentage of accidents reported per employee
- number of job hazard analyzes conducted per employee
- number of safety audits completed per year
- percentage of total operational budget allocated to safety items.

4.4 SAFETY FACTORS

The identification of leading indicators has often begun with a search for safety factors, elements or conditions that can be linked to high levels of organizational safety performance [11, 12].

Whilst there is broad general agreement about the factors that influence organizational safety [13, 14,] it is important that the specific safety factors used are appropriate for the industry. To this end, value-focused sessions were held with management from the study groups. Participants included senior management; vessel managers; safety, health and environmental management; and vetting managers. The groups’ assessments were elicited about procedures and operations in the company that could either avoid accidents or see that the correct actions were taken when exposure occurred.

The safety factors obtained were used in the case studies and refined in the light of the experience gained from running the studies. The resultant eight safety factors are:

- communication
- empowerment
- feedback
- mutual trust
- problem identification
- promotion of safety
- responsiveness
- safety awareness

These are very similar to those that the US Nuclear Regulatory Commission has recently decided to promulgate [15].

4.5 SAFETY PERFORMANCE DATA

Objective leading indicators are identified by correlating safety metrics with safety performance data. This section details the safety performance data required for the analysis. The following data is required each of the levels

Operations Data

- operational incidents frequency
- near misses frequency
- conditions of Class frequency
- port state deficiencies frequency

Health and Safety Data

- total recordable cases frequency (TRCF)
- lost time accident frequency (LTAF)
- medical treatment case frequency (MTCF)
- restricted work accident frequency (RWAF)

Similar data is collected for the business units, and/or vessel level, if those analyses are undertaken. All safety performance data requires normalization before statistical analysis to enable valid comparisons of vessels on different routes, etc. The ABS leading indicators initiative specifies how that should be done.

5. IDENTIFYING LEADING INDICATORS

Leading indicators are safety metrics that correlate with safety performance for a given organization. They can be objective or subjective measures.

5.1 OBJECTIVE LEADING INDICATORS

Objective leading indicators are identified by correlating safety metrics with safety performance data. This approach is preferred because it is objective and pragmatic. The objective leading indicators program can be done at three levels:

- organization
- business units
- fleet

5.1(a) Method Summary

The organization’s safety metrics are correlated with its safety performance data using a Spearman’s rho test. Any safety metrics that are found to be significantly correlated with any of the organization’s safety performance data are deemed to be leading indicators. The following steps are taken:

- choose safety metrics from the core metrics set and the subsidiary set
- other metrics that the organization has collected may also be suitable for inclusion
- collect safety performance data - the safety metrics and safety performance data must cover the same time period
- normalize all data
- undertake statistical analysis to ascertain which (if any) of the safety metrics are significantly correlated with the safety performance data. Spearman’s rank correlation coefficient (a non-
parametric test) is used because the data does not meet the requirements of a normal distribution.

Also, if appropriate, it is possible to investigate if the introduction of an intervention in one year correlates with a change in safety performance in the following year. In this case, the Spearman’s rho rank correlation test should be performed on each year’s safety metrics with the following year’s safety performance data. An even greater delayed effect can be investigated, for example two years’ delay, where the metrics and safety performance data are available.

5.2 SUBJECTIVE LEADING INDICATORS

Subjective leading indicators are identified by correlating survey responses with safety performance data for the previous twelve months. This section details the method employed to identify subjective leading indicators.

This method can be undertaken if the organization does not have sufficient safety metrics to look for objective leading indicators. Identification of objective leading indicators is preferred. The subjective leading indicators approach is more speculative and so should be undertaken following a survey, with the responses readily available. This approach offers the possibility of identifying new metrics for the organization to collect.

The subjective leading indicators approach uses a safety culture questionnaire developed as part of the leading indicators initiative. The safety culture questionnaire has forty Likert statements that the respondent is asked to rate on a five-point scale, with a Don’t Know option. There are five statements for each of the eight safety factors. The forty statements are divided into three sections:

- ship operations
- health and safety
- issues associated with respondent’s area of responsibility

There is also a demographics section and a small section encouraging responses and comments in free text.

5.2(a) Method Summary

- average the responses for each vessel i.e. find the arithmetic mean for the responses to the statements for all of the forty statements, for each vessel; treat missing responses as “don’t know” for up to 5% of the total responses, (where missing responses comprise more than 5% of the responses, exclude that individual’s response to that question from the analysis)
- at the same time, collect the safety performance data - one year’s data is required and this should be the most recent data available, preferably for the last twelve months, averaged to yield a single annual figure.
- normalize all data
- once all of the safety culture responses and safety performance data are prepared, begin the statistical analysis using Spearman’s rank correlation coefficient (a non-parametric test) because the data does not meet the requirements of a normal distribution
- the Spearman’s correlation analysis should be performed for each averaged vessel safety culture question response with each variable of the collected safety performance data

6. INTERPRETING THE RESULTS

The purpose of a correlation analysis is to determine if the ordering of the data (safety metrics or safety culture responses vs. safety performance data) is statistically significant. The null hypothesis states that there is no significant association of the ordering.

In the case of the objective leading indicators, the two groups are safety metrics and safety performance data. In the case of the subjective leading indicators, the two groups are responses on the safety culture survey averaged for each vessel, and each vessel’s safety performance data for the previous twelve months.

The null hypothesis for a test of correlation (here the Spearman’s rho Test) is that the two groups being investigated are not highly correlated (positively or negatively). If the null hypothesis is rejected, then there is a statistically significant correlation between the two groups. Note that an inverse (i.e. negative) correlation is of interest, i.e. as the intervention increases so safety performance measures (accidents, incidents etc.) decrease, although in the short term some interventions may lead to increased reporting.

7. UTILIZING THE RESULTS

The ABS leading indicators initiative includes a list of desired activities, attitudes and behaviors, together with a list of possible activities for improvement for each of the eight safety factors. These should be consulted when following the action plan below. The value of the leading indicators process cannot be realized until the results are incorporated into the organization’s continual improvement program.

7.1 ACTION PLAN

The following action plan should be implemented to benefit from a leading indicators assessment

- review the findings to identify the safety factors (categories of statements or metrics) that need to be addressed
• look at the appropriate safety factor which contains desired activities, attitudes, and behaviors as well as possible activities for improvement
• consider if the findings could relate to a different safety factor, as there is some overlap; in that case, consider the desired activities, attitudes, and behaviors and possible activities for improvement for that safety factor
• communicate the results to the workforce - feedback should include strengths as well as areas of weakness; this can be done in a variety of ways, e.g., written reports, team briefings
• prioritize the opportunities for improvement; initially identify three to five key areas to focus on and develop an action plan
• consider how those key areas align with other initiatives/needs
• focus on strategies that can address more than one area or need
• engage key shipboard personnel (front-line personnel) in the planning and the trialing of process changes as action plan development and implementation are typically more successful if these personnel are able to be included
• track changes for continual improvement efforts

8. FURTHER RESEARCH

The ABS Guidance Notes on Safety Culture and Leading Indicators are due to be published in 2012 [16]. The Guidance Notes are applicable to all cargo-carrying commercial vessels. The guidance helps clients to self-assess their:

- objective leading indicators
- subjective leading indicators
- safety culture

To enable clients to do this, the Guidance Notes provide full details of:

- method
- metrics tables
- safety performance datasheets
- normalization criteria
- safety culture questionnaires
- safety factors
- tips on administering the survey
- step by step guidance on statistical analysis
- worked examples
- a list of desired activities, attitudes and behaviors, together with a list of possible activities for improvement

However, even with the step-by-step guidance on statistical techniques and worked examples, ABS recognised that many clients would welcome computerised assistance with the task. To this end, ABS and Lamar University are producing a database to semi-automate the process. AP Moller-Maersk is providing the user requirements. The database is scheduled to be available early in 2012 and will be included with the Guidance Notes. The database will perform the appropriate statistical tests to complete a safety culture and/or leading indicators assessment, and provide results and recommendations in a summary report.

9. CONCLUSIONS

ABS research has developed a method for identifying potential leading indicators for improving safety performance. This research strongly suggests that it is possible to detect statistically significant correlations between some metrics (leading Indicators) and safety performance data. This is an exciting innovative approach to improving safety performance.

ABS has developed Guidance Notes on Safety Culture and Leading Indicators that are applicable to all cargo-carrying commercial vessels. These Guidance Notes enable clients to self-assess their leading indicators of safety (as well as their safety culture). Research is underway to produce a database to semi-automate the process.

10. ACKNOWLEDGEMENTS

The authors would like to thank Professor Martha Grabowski and her team for the early foundational work; and Dr. Ahmed Khago for performing the statistical analysis in the AP Moller-Maersk study.

11. REFERENCES


12. AUTHORS’ BIOGRAPHIES

Christine Tomlinson is a Staff Consultant in the Safety and Human Factors group of ABS. Dr Tomlinson has twenty five years’ experience of researching how human and organisational issues impact safety-critical systems. Since 1995 she has served on numerous safety-related committees including INTERTANKO’s Human Element in Shipping Committee, and the MCA’s Human Element Advisory Group. She is the human element advisor to the Royal Institution of Naval Architecture’s IMO Committee.

Dr. Brian Craig holds the position of Professor in the Department of Industrial Engineering at Lamar University. He teaches and performs research in the areas of Human Factors Engineering, Ergonomics, and Safety Engineering and has contracted with the ABS for the past ten years.

Maurice Meehan is the head of Sustainability and Performance for the tanker business of A.P. Moller-Maersk A/S. During this study he served as Safety Manager for the Corporate function in the Maersk Group overseeing tanker, container and supply vessels, a combined fleet of approximately 250 vessels. The focus of this position was strategic challenges on improving safety performance with a strong focus on KPIs and behaviour based safety, both shore side and at sea.