PROCESS SAFETY PERFORMANCE INDICATORS

Hans J. PASMAN

Abstract: For over 50 years to measure safety performance the Lost Time Incident Rate, LTIR was used. Fortunately, over the years the learning attitude towards accidents changed from a retrospective to a pro-active one. In the 90-s the safety management system was introduced. No management though, without the Deming cycle of Plan, Do, Check, Act, and checking, means the need of indicators. Existing LTIR-values were used not realizing these reflect personal rather than process safety. In 2005 after the BP Texas City refinery vapor cloud explosion, awareness of the difference broke through and Process Safety Leading and Lagging Metrics were formulated. In January 2012 an international conference was held in Brussels organized by EPSC and CEFIC. Results will be summarized. The paper will explain briefly, where we are now, and what still is ahead.

Key words: Process safety, Management indicators, Bayesian networks, Risk assessment.

Introduction

One of the first issues to discuss in the newly founded Working Party on Loss Prevention and Safety Promotion in the Process Industries (as its full name was) of the European Federation of Chemical Engineering (EFCE) in the first half of the 1970s was the cause of the order of magnitude gap in an indicator value between American companies and European ones. This concerned the LTIR - Lost Time Injury Rate - or LTIF - Lost Time Injury Frequency - as hours lost per unit of total worked hours in a plant. As registration unit different numbers of hours were in use, but most often the rate was expressed per one million worked hours. At the time we didn’t take the step to generalize it as an indicator, but just discussed the causes of the differences of this rate since it was the only statistic available. European plants had roughly a factor ten higher rates than the U.S. ones. All kinds of bias possibilities and causes were mentioned. Registration biases could, however, at a certain stage is ruled out because the BASF plant in the U.S. had also a ten times lower rate than the plants in Ludwigshafen. Since it further appeared that each year considerable improvement was made, as shown for example by the Shell company record in Fig. 1, which is typical for the branch, after a while nobody asked for the background anymore.

Improvement of the LTIR is still continuing. Fig. 2 gives a more recent picture of the International Association of Oil & Gas Producers and of the Dow Chemical Company.

![Graph showing LTIF from 1965 to 1990 for Contractor and Company staff]

Fig. 1 Example of improvement made over the years in reducing the number of accidents with the effect of an injury preventing the worker to be at work for some time for a major company as Shell and how smaller firms picking up outsourced (maintenance) work, had to catch up with the trend (Visser, 1995)

In particular at the end of the 1990s there was a strong optimism that the state of zero accidents could be reached soon, as illustrated by the graph from the Norsk Hydro Annual report 1998.

Beside the trend of preventing accidents, it was also the time of increasing economic pressure, downsizing of personnel, early retirements, less solid education, job hopping, saving on maintenance, more complex plants, etc. which all together countered the trend of increasing process safety. Plants run

---

1 Texas A&M University, Artie McFerrin Department of Chemical Engineering, Mary Kay O’Connor Process Safety Center, College Station, USA, hjpasman@gmail.com
by people with less overview and insight of what can go wrong, who are less alert due to all kinds of distraction, or who violate rules because it saves time and experienced nothing serious happened, can suddenly be confronted with an accident. In general, these latent threats went unnoticed at board level, because the LTIR was still further decreasing and safety was therefore not a priority.

Kletz (Kletz, 1993) had warned for such situation many years ago and Körvers (Körvers, 2004) motivated his promotion work on accident precursors in 2000 - 2003 with references to Hopkins and Hale (Hopkins, 2000; Hale et al., 1998; see also Sonnemans and Körvers, 2006). Hopkins mentioned the Longford gas plant in Australia where in 1998 a disastrous accident occurred while the LTIR of that plant was exemplary and Hale wrote that formulating reliable safety performance indicators constitutes a fundamental problem in safety science. It took till the 2005 BP Texas City refinery vapor cloud explosion (CSB, 2005) that - with a shock - the distinction between personal safety with the LTIR as indicator and the more hidden process safety became clear and installing process safety performance indicators became a priority item. The report by the Baker Panel (Baker et al., 2007) following the CSB investigation of the accident and examining all BP’s refinery activities in the U.S. uncovered the leadership and management weaknesses and safety culture issues.

In fact, Patrick Körvers in his dissertation (Körvers, 2004) investigated many cases: first using an accident data bank and later by doing investigative work in more detail at three process plants. He was looking to find (hidden) accident precursors by checking process disturbances or deviations. If these reproduced and appeared to be safety critical, their cause was investigated applying a three layer hierarchical system of operational level, tactical, and strategic level; see Fig. 4. Each layer can be thought of as a control system consisting of a sensor
(observation), a processor (judgment) and an actuator (intervention), controlled by a supervising element of the next higher layer. Of each overlooked precursor the causation was traced and the failing element identified. It turned out that the majority of unnoticed precursors could be ascribed to management failure.

**Materials and methods**

*Leadership, management oversight, and control*

As was realized already for a long time, in safety the role of leadership and good management is crucial. Since the Piper-Alpha disaster in 1988 and the push given by the Lord Cullen report the process industry introduced in the 1990s safety management systems. However, in general no explicit provisions were taken to check the effectiveness of the measures. A number of routines, such as having written procedures available, performance of HazOps and LOPAs, management of change, audits, training of personnel, etc. had been installed, but when year after year no near misses or worse happen, motivation to keep such routines upfront in mind can easily erode.

Management takes decision on the basis of indicators. As regards financial indicators (profit, losses) that is clear, but Deming introduced in the early 1950s the ‘wheel’, see Fig. 5, or the Plan, Do, Check and Act cycle for product quality, where the Check-stage can be interpreted as obtaining indicator values upon which corrective action can take place. So, early in the 2000-time frame, it was ‘in the air’ that safety performance indicators were necessary. The Working Group on Chemical Accidents of the OECD (Organization for Cooperation and Development, with head office in Paris) issued in 2003 an interim Guidance on Safety Performance Indicators (OECD, 2003) to support initiatives to establish indicators as e.g., the Responsible Care program of the chemical industry. UK Health and Safety Executive (HSE, 2005) produced a practical guide. This was followed after the Texas City explosion at the BP site by publications of the Center for Chemical Process Safety (CCPS, 2007a and b, 2010). Key Performance Indicators (KPIs) or Process Safety Performance Indicators (PSPIs, PSIs) suddenly got much attention, as appeared from various publications, such as of the UK Oil and Gas industry; the Step Change in Safety (UKOOA, 2006a and b), the American Petroleum Institute recommended practice 754 (API, 2010), and the guideline by the European association of chemical industries (CEFIC, 2011). The European Process Safety Centre, EPSC and CEFIC organized last January 2012 an international conference on the topic.

Indicators can be divided into two main groups, the lagging and the leading ones defined as follows:

**“Lagging” Metrics** - a retrospective set of metrics that are based on incidents that meet the threshold of severity that should be reported as part of the industry-wide process safety metric.

**“Leading” Metrics** - a forward looking set of metrics that indicate the performance of the key work processes, operating discipline, or layers of protection that prevent incidents.

**“Near Miss”** and other internal Lagging Metrics - the description of less severe incidents (i.e., below the threshold for inclusion in the industry lagging metric), or unsafe conditions, which activated one or more layers of protection. Although these events are actual events (hence, lagging), they are generally considered to be a good indicator of conditions that could ultimately lead to a severe incident (CCPS, 2007a) and as such leading metric.

Management takes decision on the basis of indicators. As regards financial indicators (profit, losses) that is clear, but Deming introduced in the early 1950s the ‘wheel’, see Fig. 5, or the Plan, Do, Check and Act cycle for product quality, where the Check-stage can be interpreted as obtaining indicator values upon which corrective action can take place. So, early in the 2000-time frame, it was ‘in the air’ that safety performance indicators were necessary. The Working Group on Chemical Accidents of the OECD (Organization for Cooperation and Development, with head office in Paris) issued in 2003 an interim Guidance on Safety Performance Indicators (OECD, 2003) to support initiatives to establish indicators as e.g., the Responsible Care program of the chemical industry. UK Health and Safety Executive (HSE, 2005) produced a practical guide. This was followed after the Texas City explosion at the BP site by publications of the Center for Chemical Process Safety (CCPS, 2007a and b, 2010). Key Performance Indicators (KPIs) or Process Safety Performance Indicators (PSPIs, PSIs) suddenly got much attention, as appeared from various publications, such as of the UK Oil and Gas industry; the Step Change in Safety (UKOOA, 2006a and b), the American Petroleum Institute recommended practice 754 (API, 2010), and the guideline by the European association of chemical industries (CEFIC, 2011). The European Process Safety Centre, EPSC and CEFIC organized last January 2012 an international conference on the topic.

Indicators can be divided into two main groups, the lagging and the leading ones defined as follows:

**“Lagging” Metrics** - a retrospective set of metrics that are based on incidents that meet the threshold of severity that should be reported as part of the industry-wide process safety metric.

**“Leading” Metrics** - a forward looking set of metrics that indicate the performance of the key work processes, operating discipline, or layers of protection that prevent incidents.

**“Near Miss”** and other internal Lagging Metrics - the description of less severe incidents (i.e., below the threshold for inclusion in the industry lagging metric), or unsafe conditions, which activated one or more layers of protection. Although these events are actual events (hence, lagging), they are generally considered to be a good indicator of conditions that could ultimately lead to a severe incident (CCPS, 2007a) and as such leading metric.
At the EPSC-CEFIC conference 31 January and 1 February 2012 Executive Director of CEFIC, Mr William Garcia, showed the indicator categories as in Fig. 6. It was also advised to not discuss whether an indicator is ‘lag’ or ‘lead’ for a time period longer than 10 minutes! In Safety Science journal Volume 47 in 2009 a discussion has been held at academic level triggered by an ‘Introductory story’ by Hopkins (Hopkins, 2009). A practical analysis was by Hudson (Hudson, 2009), who suggested considering control theory. Lagging indicators are to be compared with negative feedback and as such relatively simple: correct and recover. Leading indicators on the other hand mean feed-forward control which requires in fact a model to establish what correction and how much correction should be given. As a model a bow-tie (fault tree connected to an event tree of consequences by a central hazardous material release event, also called initiating event) could serve because it depicts the scenarios possible for an installation leading to breaches of process safety which can be quantified as risks. This implies, however, that various ‘soft’ organizational and management influences and human error must be included in this model as well. Although this will generate new problems, it is a direction which should be further explored and to which we will return.

The conference consisted of four working sessions between an introductory and a summarizing plenary. In the Introductory plenary Mr. William Garcia, mentioned above, spoke on behalf of the ICCA, the International Council of Chemical Associations, promoting the introduction of indicators in the companies globally. Mr Kenan Stevick of Dow Chemical presented the position of the American Chemical Council (the equivalent of CEFIC) which is supporting a single global set of lagging process safety metrics, while for leading indicators flexibility is needed and companies should choose their own. Incident severity shall be component of any metric. The metric scheme should become an ISO standard. The subjects of the four sessions in break-out groups were:

- **Session 1:** Implementing process safety performance indicators. This consisted of a series of industry speakers explaining their experiences to encourage others to start the process. Once a system runs, 3-5 years are required to get it stable.

- **Session 2:** Broaden the basis, was meant to discuss ways to convince the smaller companies to introduce indicators. Smaller companies have the handicap of possessing in general less expertise, sometimes even no safety management system. If they get indicators the results do not show a neat trend line but will fluctuate. On the other hand, there is enough to gain, also cost-wise, if they would control their operation better.

- **Session 3:** Going public. This final goal is in line with that of Responsible Care to get the public to trust the industry more. Pros and cons of opening the indicator results to the public were discussed. General conclusion was to make them public (which is also in line with the proposed Seveso 3 Directive), but only after sufficient experience has been obtained.

- **Session 4:** Leading indicators: Three main groups of leading indicators can be distinguished:
  - Mechanical integrity indicators (inspections, controls),
  - Action items follow-ups: (PHA - Process Hazard Analysis, audit and near miss actions),
  - Training/competence indicators in the form of quality testing, such as percent of personnel trained and how complete roles in process safety are defined and assigned.

Companies can select leading indicators according to their needs.

As follow-on, in June a meeting has been held in Paris between representatives of OECD and CEOs of companies to discuss establishing process safety governance.

**Results**

In fact, the CCPS guidelines on the topic of process safety performance indicators resulted in a number of close to 400 possible indicators measuring the effectiveness of 22 management system elements. With so many indicators overview will be lost: only 5 to maximum 10 indicators can be handled sensibly. On the other hand, in safety ‘the devil is in the detail’ and one has to monitor the right details to find out where risk can develop. The solution to this dilemma will be aggregation. Hassan and Khan (Hassan and Khan, 2012) worked out an example of aggregation of a hundred specific indicators via some 40 key asset integrity indicators, 13 activity indicators to 3 element indicators for comparing safety performance of five oil and gas plants. In Fig. 7 the levels of aggregation are shown. The aggregation is risk-based. It means that, first of all, the importance of each indicator for the overall safety and well-functioning of the plant is obtained as a weight factor by expert opinion in an AHP (Analytic Hierarchy Process) through a pair-wise comparison matrix. Secondly, the result of a lagging indicator is expressed as the frequency of the occurrence of...
an incident multiplied with its severity, and that of a leading indicator as the percentage of the success multiplied with the importance of success. The result was quite satisfactory.

Fig. 7 Aggregation levels of indicators according to Hassan and Khan (Hassan and Khan, 2012), to provide insight through key indicators at work floor supervision level, activity ones to serve middle management, and three element indicators: operational integrity, maintenance integrity and personnel integrity for monitoring by top management.

As long as the indicator values improve and there are no large fluctuations in annual results, there will be no need to have an additional criterion for what is acceptable or not. This, however, will change as soon as values become stagnant or worse, when they decline. In that case a more direct relation with the magnitude of risk will become desirable. Given a risk expressed as a product of frequency and a severity, various options for decision making are available: use of a semi-quantitative risk matrix as e.g., in the standard IEC 61508 (IEC 61508, 2010), comparison with business risks expressed as e.g., in the standard IEC 61508 (IEC 61508, 2010), comparison with business risks expressed in monetary units, legal constraints, and insurance requirements. Now we return to the bow-tie concept mentioned by Hudson (Hudson, 2009).

Bow-ties are composed of a fault tree and an event tree part. Both can to a certain extent be quantified, and they can be very helpful in determining the effectiveness of preventive and protective barriers. However, the reliability of human action has never been incorporated in a detailed way because of its complexity. This is slowly changing. Thanks to the Bayesian Belief Network technique allowing larger flexibility and versatility in probabilistic modeling, more complex causal relationships can be simulated. Causality has been studied by pioneers as e.g., (Pearl, 1988; Morgan and Pearl, 2000), in the artificial intelligence domain to enable machine decision making. As a result of practical software becoming available that takes away the computational burden, it has the last few years found very useful applications in medical science, social sciences, economy and finance, and software reliability determination. Recently it has been applied in aviation safety (Mohaghegh et al., 2009; Groth et al., 2010; Ale et al, 2009) and is starting to find its way in Quantitative Risk Analysis of processes in on- and offshore (Khakzad et al., 2011; Pasman and Rogers, 2012a a,b; Vinnem et al., 2012).

Bayesian Networks (BNs) are cause-consequence chains and consist of nodes (also called vertices) and arcs (or edges). A node represents a stochastic variable, an arc a dependency. The simplest network is formed by two nodes, $X$ (cause) and $Y$ (effect), connected by a directed arc (an arrow from $X$ to $Y$ reflecting the dependency). Networks are acyclic (an effect cannot be the cause of its cause!) and the variables can be discrete or continuous. In case of discretely valued variables with a certain probability to be in the states true or false, or, working or failed, the arc implies in the dependent node a conditional probability, the value of $Pr(Y|X)$ or the probability that $Y$ gets in a failed state given $X$ fails, hence the strength of the influence of $X$ on $Y$. In a multi-node network this will take the form of a conditional probability table. Solution is exact but gets cumbersome to determine if the nodes become numerous. In the continuous network a variable can have the form of a continuous probability distribution function, and the relationship can be expressed arithmetically, and solution to derive the joint distribution is by sampling. The networks are called Bayesian because the Bayesian theorem of availability of new evidence enabling updating and inference forms the core of its success. The networks are useful as tool to predict given root causes the probability of effects, but also the other way around; i.e., to diagnose causes on the basis of observation of consequences. Because of the network’s probabilistic nature, a node can be given a value on the basis of a belief (expert opinion) and the dependence between two nodes need not to be deterministic and not even to be a correlation, but can be just a rank correlation. The rank correlation is the product moment correlation of the ranks of variables $X$ and $Y$, and measures strength of the monotonic relationship between the two and hence the influence of one on the other. In particular a non-parametric continuous network, such as Uninet of Delft University of Technology, is suitable for that purpose (see e.g., Morales et al., 2008). All bowties can be modeled as a BN with much more possibilities to incorporate the complexities encountered in practice such as multi-mode and time effects, than the conventional techniques.
The aviation safety application mentioned above went farthest in modeling human action and decision making in an organizational framework including error probability resulting in a BN of more than a thousand nodes. Basis for this was also the work by Bellamy on occupational safety and occupational accident modeling, described in e.g., (Bellamy et al., 2008; Ale et al., 2008). They applied accident data to create scenarios (Story-builder) schematized those in bow-ties including barriers.

Shown is a bow-tie of the first gas-oil separator of a processing unit of an offshore platform according to a description and reliability data of Khan (Khan et al., 2002) and modeled as a Bayesian network. The separator contains about 5000 kg of hydrocarbons which by overpressure can BLEVE or as a leak can form an explosive vapor cloud. The consequences part to the right of the critical event is expanded to calculate the damage cost due to loss of life and materiel. Added are the three top asset integrity indicators to the left with the weight factors defined by Hassan and Khan (Hassan and Khan, 2012), where the personnel indicator is taken as the dominant one. The indicators influence parts of the installation of which can be assumed that their functioning is much dependent on the quality of operation and maintenance or on right design (e.g., component no. 12: safety valve designed too small). The influence is constituted such that the failure rate is inversely proportional with the integrity value.

A change of a factor of two in personnel integrity implies in this set-up a change of a factor of ten in final risk as shown in the table at the bottom right.

Vinnem et al. (Vinnem et al., 2012) developed after several preparatory studies, following Ale et al. and Groth (Ale et al., 2009; Groth et al., 2010), and further building on an extensive experience with offshore technical risk assessment, a model integrating organizational, human and technical factors. To that end the authors introduced the concept of risk influencing factor or RIF. RIFs come in two layers: bottom layer are management properties influencing the upper layer of work force conditions, such as competence, available documentation, and time pressure. The upper layer of RIFs influences probability of human mistake, of slips and lapses and of acts of violation.

From the approach by Vinnem et al. (Vinnem et al., 2012) to a coupling with indicator values which are exponents of the safety management system is not a big step. In fact, Pasman and Rogers, building on an earlier paper of Knegtering and Pasman (Knegtering and Pasman, 2012), are presenting a paper at the MKOPSC symposium, October 2012 about the principle as illustrated in Fig. 8. The approach is to a certain extent different from the one by Vinnem et al. (Vinnem et al., 2012), in the sense that risk factors are defined both from organizational and human nature as well as from environmental and technical origin. Indicators are considered as
just part of the risk factors with the leading ones mostly influencing safety on the long or middle term, while other risk factors, such as equipment vibration, bad weather and many others, may act on the short term. The idea is that the model may produce also an urgency of taking measure. All this is rather preliminary and has still to be worked out in concrete projects.

Conclusion

• Indicators are indispensable to monitor the state of process safety in a plant. Lagging indicators are based on incidents. The threshold of severity for counting them as significant shall be standardized. Leading indicators shall be selected freely. Near misses are important and shall be meticulously investigated and corrections implemented.

• International/global acceptability of a standard on indicators is sought. As soon as indicator results are stable and can be relied on, which can take 3 - 5 years, it is desirable to go public to gain trust.

• As long as year after year improvement can be shown, there may be no need for a criterion of acceptability, but if results become stagnant or even get worse, a criterion will become required. Connecting indicators with a risk level and assessing the risk in a risk matrix will be a way to proceed.

• Modeling of risk with inclusion of organizational, management and human factors with the potential use of indicator values is at an initial stage. Progress is being made due to the advent of Bayesian Belief network software allowing complex causality relations in probabilistic terms, so that ‘vague’ influences and expert opinion can be incorporated.

References


