SAFETY METRICS
Corporate & Site-Level Scorecards
By Paul A. Esposito

Metrics: what you measure happens. For more than 50 years now, W. Edwards Deming and others have extolled the value of upstream and process metrics as a means of driving bottom-line results (e.g., outcome metrics such as incident rates).

The well-known quote attributed to Deming says, “If you measure the process, results will come.” Leading companies recognize that safety performance is now defined by a balanced scorecard, where multiple financial, process, learning and growth, and customer metrics are combined to better measure both the progress and status of a company or program (Kaplan & Norton, 1996).

Petersen (2001) published a list of six essential and suggested metrics. This article explores leading metrics to measure safety performance presented in a balanced scorecard format, discusses what metrics may be most valuable to start with and identifies risk-centric alternatives to incident rate metrics.

Key Financial Metrics
• Workers’ compensation: While not consistent across states or countries, it is one measure to keep track and compare annually.
• Budget implementation: Track the budget for safety improvements throughout the year. If budget is taken away in the third or fourth quarter, this should impact the overall safety performance and be reported.
• Manpower: In addition to other resources, this includes things such as time for employees to participate in safety committees, conduct inspections, staffing and available time, and access to third-party consultants.

Key Process Metrics
• Safety management systems (SMS) assessment scores: Leading companies such as GE, Honeywell, Standard Aero and Eaton use the key element audit management systems as a means of measuring process. This can be seen in ANSI Z10, ISO 45001, 14001 and 9000, and OHSAS 18001. If the audit process is sound and substantial in its verification of process (not just documentation), it is valuable not only as a means of reporting accountability and process improvements, but also as a tool to help operational organizations know where to improve next.
• Management systems improvement initiatives: Assessments typically generate improvement initiatives. The timely closure of these initiatives is also tracked as part of a global scorecard.
• Risk reductions: Moving from an incident-based to a risk-based approach to metrics is also more in line with the current trend toward world class. Risk reductions are measured by many different means.

Depending on the data available, some measure: Number of residual risks and their trend based on initiation of safety improvements; upstream data or the number of corrective actions implemented that had elimination, substitution or engineering controls; closure rate of these corrective actions; or the reduction in people required to wear PPE.

A subset of risk reductions is to measure the specific reductions in risk factors, such as severity or consequence levels and likelihood levels.

Another subset is the hierarchy of controls. Adding additional controls and additional layers of controls is how risks are reduced. Only new or additional controls or mitigations reduce severity or likelihood. So, leading companies measure the number of hazards eliminated, the number of new controls, the number of new engineering controls and the number of tasks that now have multiple layers of controls.

• Critical to safety failure rates: Many of the better engineering controls require preventive maintenance to maintain their readiness. Controls such as cranes, hoists and fire protection systems are critical to work as designed. Often, third parties maintain these for a facility. Failure rates of these critical components should not be a secret tucked inside of the preventive maintenance database, but should be included on a regular monthly report that the testing was competed and that there were no failures. If failures occur, preventive maintenance schedules may need to be revised and other corrective actions put in place.

Learning & Growth Metrics
• Training completion and retention: Training completion by itself is a lagging metric. It does not factor in quality of the instructor or learning that occurred. Often, as part of an inspection or observation program, knowledge-based quizzes are incorporated to generate a retention metric. Many other retention metrics can be implemented that are often added to the scorecard, such as instructor evaluations or specific exam question failure rates.

• Conformance rate: Significant or critical to safety controls based on inspection and observation data. Significant risks and their controls are identified. The conformance of these specific controls is calculated as a conformance rate as a result of site inspection and observation programs. Typically targets of 95% to 100% are expected.

• At-risk trend reductions: When trends occur for inspection or observation nonconformances or at-risk behaviors, Pareto analysis determines the most serious and pervasive trends. These trends are then subject to root-cause analysis and subsequent preventive actions. The number of repeat findings could also be trended.

• Incident hazard trends: When performing incident analysis, the hazardous energy that was supposed to be controlled can be listed and trended. Again, any significant trends can be an opportunity to improve the risk assessment process.

• Incident hierarchy of control trends: Trending the highest level of control failure provides valuable insight into what is working and not working, especially for incidents that have occurred. Prevention plans help reduce these trends.

• Incident causal factors: Pareto analysis of the causal or contributing factors is especially valuable if trends are apparent, and corrective actions can reduce these trends. Corrective actions may or may not
address the real issues, trends may not recognize any significant change. Monitoring these trends helps verify the efficacy of corrective actions.

- Incident root causes: If the root causes turn out to be safety management systems, Pareto analysis can be especially helpful. In addition to an SMS audit, this is another way to validate the strengths and weaknesses of management systems.
- Compliance: This is an optional leading metric typically generated from some type of facility audit and performed by the site with third-party audit verification. Often, compliance will be verified for significant high-risk exposures. Regulation citations or violations are often recorded within this metric.

Customer Metrics
- Total incident rate: This is a lagging metric of injury frequency.
- Days away and restricted case rate: This is a lagging metric of severity. These are the two significant metrics of OSHA’s Voluntary Protection Programs, along with the SMS audit and closure rate of SMS action plans.

- Serious injuries and fatalities: When potentially catastrophic actual or near-hit incidents are recorded they are calculated, tracked and trended.
- Number of overexposures: Occupational illnesses can often take years to manifest themselves due to the latency period of many of the more potent substances and physical hazards. In advance of the manifestation of illness, overexposure reductions can help prevent occupational illness if targeted and measured (AIHA, 2001).
- Employee engagement: Many different means can be used to measure employee engagement. Some companies measure the number of suggestions or positive recognitions awarded. In other cases, companies measure the percentage of people who get involved in various safety activities such as inspections. A percentage is a better measure of the breadth of involvement, as sometimes the involvement can come from the same small group of employees.
- Perception or culture surveys: Surveys are a valuable tool for determining the gap between the perceptions of workers and those of management. Real or perceived, perception and cultural differences in

Celebrating 25 YEARS
OSHA®
Training Institute Education Centers

Visit us at Booth #2830

osha.gov/otiec

EXPERIENCE THE DIFFERENCE
OSHA Training Institute Education Centers provide safety & health education nationwide for the benefit of public & private sector workers and employers.
communication and expectations or accountabilities are a continual process that many companies try to improve each year. Categories are scored on the Likert scale, while online tools make designing and delivering these surveys easier. One weakness is the delay between the survey and getting improvement initiatives in place. The longer the time lag, the less effective the next survey will be as participants may suspect management does not value their opinions.

While many of these metrics take time to develop, an audit program is a common starting point. The audit could be a first data collection mechanism to begin verifying some of the metrics. The audit is also normally completed by an independent third party, at least as part of the audit team, to lend some objectivity to the generation of metrics. The audit can generate some of the other metrics (e.g., systems scores, closure rate of management systems improvement initiatives, employee survey summaries).

**Site/Organizational Level**

At the site or organizational level, an additional level of detail is typically needed to better monitor, review and react to the various safety programs and initiatives. Specifically, the amount of change and closure are the key data points, along with some measure of quality. The lower into the organization you can measure, the more accountability you have.

Like any situation in life, you need the right tool for the job. Risk assessment yields a full breadth of metrics. Initially the number of assessments can be a target, but to be worth the time investment, changes must be identified, otherwise it is simply a paper exercise. These changes must be closed and stay closed, so some organizations add a 30-day timer to the closure as an additional metric. Finally, effectiveness is based on the type of control that gets implemented. More effective controls include substitution and elimination or engineering controls, rather than just training or PPE. Whether the target is a number or a percentage defined or compared to the previous reporting period, comparison against a target is essential.

Building up to this level of scorecard may take time. Thus, the circled metrics in Figure 1 are common starting points. In reality, the starting points may actually be the data you have, rather than what you want, until the data you want is available for collection.

Big data is a buzz phrase now making its way into the safety and health vocabulary. Many software providers offer platforms for leading companies to manage and analyze the myriad of potential safety performance data.

**Metrics of Little Value**

This article does not discuss some common historical safety metrics. These metrics have often been misapplied, misunderstood or overvalued. Some safety professionals may take a different view; any diversity in the debate or discussion can only help strengthen the use of metrics in safety and health programs. For example:

- Injury by parts of the body: If PPE is the ultimate solution, this may be an effective metric. Otherwise, it is addressing only a symptom.
- Incidents by time of day: There is no argument that when shifts run into extra hours, the frequency of incidents increase. Nighttime versus daytime activities are also viable. Fatigue is absolutely a confounding factor. However, the morning, lunch and evening analyses are only partially viable if administrative controls such as training or awareness are the ultimate risk reduction weapon. This type of thinking also leads to blaming employees.
- Unsafe acts versus unsafe conditions: Petersen has always maintained that these are just symptoms of something wrong with the management system. Who is responsible for the unsafe conditions? Have any of these been tied to management decisions not to resource higher-level controls? Workers do not create the hazards in the workplace. They can only strive to implement the control given to them, effective or not.

**FIGURE 1**

**DEPARTMENT SCORECARD**

<table>
<thead>
<tr>
<th>OSH preventive maintenance report topics</th>
<th>No. of events</th>
<th>No. of changes identified/ needed</th>
<th>Closure rate (from date closed)</th>
<th>Effectiveness (control type or conformance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current program risk assessment changes to controls</td>
<td>Actual</td>
<td>Target</td>
<td>Actual</td>
<td>Target</td>
</tr>
<tr>
<td>Monthly safety review incidents</td>
<td>5</td>
<td>20</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Monthly inspections or discrepancies high-risk control conformance</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Monthly observations conformance or % safe operations</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Supervisor communication (OSH toolbox) resulting in suggestions</td>
<td>50</td>
<td>10</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>OSH action plan status</td>
<td>10</td>
<td>20</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>OSH rewards and recognition</td>
<td>50</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• Number of observations: Observations alone do not necessarily change behavior. The ABCs of behavior motivation, attitudes and beliefs are dictated more by the volume of antecedents (A) and positive consequences (C). More astute organizations set goals for the number of new ideas generated by the observation dialogue process. Closure of these ideas and suggestions are the positive consequences that drive behaviors.

• Zero incidents: Without a doubt, this is the outcome expected from all of our efforts. But as a metric it is lagging and as a measure it is only an indicator that must be measured and reported. It is important, however, as the primary metric or goal it is akin to driving a car by only looking in the rearview mirror. It does not measure safe. Risk reduction, the number of new controls, the elimination of hazards and conformance rates measure safe. Zero nonconformances would be a much improved metric.

Like any situation in life, you need the right tool for the job. Metrics can be used to motivate or demotivate depending upon delivery, methodology and application.

**Conclusion**

Leading organizations understand that there is no single metric for either the entire program or the individual component parts. A balance, or complementary set of leading and lagging or, better input, process and output metrics are used in addition to the standard lagging metrics such as workers’ compensation and incident rates to drive continual improvement in safety performance.

These metrics are just the beginning of the list. They are the new measure of safety performance. Line management is held accountable to these metrics as part of the performance appraisal process. At the worker level, these metrics become the basis for recognition programs. There is no substitute for holding the right people accountable to the right things. In addition, if there is an overwhelming recognition to the positive rather than negative metrics, people are motivated to provide discretionary effort. Discretionary effort is the one cultural dynamic that can give any company a competitive advantage.

**References**


Paul A. Esposito, CSP, CIH, SMS, is president of STAR Consultants Inc., a business that provides advanced level strategic planning, training and systems development consultation to Fortune 500 and leading organizations worldwide toward the continual improvement of their safety and health programs. Esposito holds an M.S. in Safety and Health from Johns Hopkins University and is a professional member of ASSP’s Chesapeake Chapter which he serves as a Delegate and Nominations and Elections Chair.