



## On the use of criteria based on the SMART acronym to assess quality of performance indicators for safety management in process industries

Jon Tømmerås Selvik<sup>a,b,\*</sup>, Surbhi Bansal<sup>a</sup>, Eirik Bjorheim Abrahamsen<sup>a</sup>

<sup>a</sup> University of Stavanger, 4036, Stavanger, Norway

<sup>b</sup> NORCE Norwegian Research Centre AS, 4068, Stavanger, Norway

### ARTICLE INFO

#### Keywords:

Performance indicators  
Safety  
Barrier management  
SMART  
Criteria  
Quality  
Process industries

### ABSTRACT

Management of safety, and barriers in particular, includes using information expressing performance, i.e. use of safety performance indicators. For this information to be useful, the indicators should demonstrate adequate quality. In other words, they should satisfy some predefined set of quality criteria. Without showing adequate quality, the indicators are generally unable to provide sufficient support for barrier management, which could result in poor decisions. In this article, the use of the SMART criteria is considered to assess the quality of safety performance indicators in process industries. SMART being an acronym for 'specificity', 'measurability' or 'manageability', 'achievability', 'relevancy' and 'time-based', covering five key aspects and criteria for assessing the quality of an indicator. A discussion on whether the indicators are able to demonstrate adequate quality by satisfying these criteria has been conducted. The finding is that all of the SMART criteria should be satisfied for a safety performance indicator to demonstrate acceptable quality and to be regarded as useful to support barrier management decision-making. However, it has also been observed that including the 'M' criterion in the assessment of quality is not needed. When all the other criteria are satisfied there is no way the conclusions could be misleading as a result of measurability or manageability aspects. Hence, for safety performance indicator quality, only four of the criteria are assessed and suggested for such situations to shorten the acronym to 'STAR'. A key safety indicator used in downstream process facilities, i.e. 'dangerous fluid overfilling events', motivated from the 2005 Texas City refinery accident, is used to illustrate the situation. The indicator is also applied to another incident, the Buncefield oil storage depot's accident in 2005, to provide a broader context for using it. The findings in this article could also be applied beyond the context studied. This means that, despite focusing on safety indicators in the process industries, the findings are considered as relevant and applicable to other types of performance indicators and to other energy industries.

### 1. Introduction

In this article, the focus is on achieving useful performance indicators to support decision-making related to safety and barrier management in the process industries. For example, when adopting the "safety diagnosable principle" or "defence in depth" it is essential to have appropriate indicators measuring barrier conditions; see Saleh et al. (2014a; 2014b). A variety of safety performance indicators (SPI) are used for this purpose and included in indicator portfolios to provide a sufficiently broad information basis. However, the usefulness is challenged by quality, as information from some indicators might be misleading or totally disregarded in practise but nevertheless be associated with costs. Consequently, assessment of SPI quality is an important activity related

to the construction and use of the performance indicator portfolio. Adequate quality links to the ability to meet safety target and business goals, and visions.

One common and in principle simple way to assess the quality of performance indicators is by using the SMART criteria, referring to five standard criteria covering main quality aspects (Badawy et al., 2016; Parida and Kumar 2006; Doran 1981). Basically, by verifying that the indicators satisfy the criteria, one avoids spending resources on collecting and analysing information not contributing with any or with poor business value. SMART being an acronym for:

- Specificity
- Measurability

\* Corresponding author. University of Stavanger, 4036 Stavanger, Norway.

E-mail addresses: [jon.t.selvik@uis.no](mailto:jon.t.selvik@uis.no), [jssel@norceresearch.no](mailto:jssel@norceresearch.no) (J.T. Selvik).

<https://doi.org/10.1016/j.jlp.2021.104392>

Received 12 October 2020; Received in revised form 1 December 2020; Accepted 8 January 2021

Available online 22 January 2021

0950-4230/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

- Achievability
- Relevancy
- Timeliness

These are further described in Section 3 and in Table 1.

Despite being commonly used, and quite intuitive in their relation to assessment of quality, it is not obvious that these criteria meet the objective of demonstrating SPIs with high quality, despite there being extensive literature available on different benefits and challenges related to performance indicators. In this article we focus on the SPI quality and relation to the SMART criteria, aiming to provide some clarification regarding how suited the SMART criteria are for the safety and barrier context. For this, we question whether these five criteria are appropriate for assessment of quality, or whether some adjustments are called for. There could be a need to add other criteria or reject some of those already present. Relevant criteria could perhaps be left out due to poor incentives, for example to keep the nice acronym created or simply be context related.

In a previous study by Selvik et al. (2020) discussing the use of these criteria in a general business context, it is suggested an ‘M’ swap, i.e. to include an assessment of ‘manageability’ instead of ‘measureability’. This latter criterion is considered to make more sense when dealing with key performance indicators compared with business goals. Making the swap should make the SMART criteria better suited for assessing quality. See also discussion in Section 4. However, this is not necessarily the situation when studying indicators in a safety context, as there could then be other quality aspects being relevant. Particularly, it is not obvious that a ‘manageability’ criterion is needed, as there should in principle always be possible to perform some safety-related action to improve current situation, otherwise it challenges the need of the ‘relevancy’.

Regarding the assessment of SPI quality in the process industries, we believe it is important to consider the appropriateness of the SMART criteria as basis for demonstrating SPI quality. An objective of the article is thus to contribute to an improved framework for performing the assessments. As a basis for the discussions, we include also consideration of other criteria that could be applicable for the assessment of quality, being suggested in literature, such as e.g. adding ‘explainability’ and ‘relativity’ to extend the acronym into ‘SMARTER’ (Better Regulation Task Force 2000). There are also several other alternatives as presented by the overview in Section 3.

The article is structured as follows. Section 2 gives a brief introduction to barrier management and use of safety performance

**Table 1**  
SMART criteria for assessment of performance indicator quality.

Criterion	Description
<u>Specificity</u>	Precision; the indicator should be sufficiently precise. It should be clear what the indicator expresses (measures); the parameters of the measure should be unambiguous; and the numbers should not depend on who is producing them and who is interpreting them (i.e. consistent interpretation).
<u>Measurability</u>	Comparability; it should be possible to quantify and compare to other data, e.g. progress towards the attainment of the objectives, where it should reflect the level of general development in a certain aspect. The data on the parameters defining the indicator measure should be collectable and available in sufficiently high quality.
<u>Achievability</u>	Attainability; it should be possible (realistic) to achieve the objectives on which the indicator is based. The indicator should provide adequate information, with respect to confirming attainment of the objective.
<u>Relevancy</u>	The indicator should provide essential information for business management and improvement (i.e. aligned with business objectives). The indicator should thus be important for business performance.
<u>Time-based</u>	The indicator value should cover an appropriate period (a predefined and relevant time-frame period). Too short a period provides limited knowledge about the aspects studied.

indicators, where different types of indicators can be combined into portfolios. Then Section 3 summarises the five SMART criteria. In Section 4, we discuss whether these five criteria in themselves are appropriate to use for the assessment of quality for a selected performance indicator. We also point to other criteria suggested in literature that could be considered. Then, in Section 5, we discuss how to combine the individual SPIs into a portfolio useful for decision-making purposes. In Section 6, we consider the overall perspective and discuss the use of the SMART criteria from a portfolio perspective, and how the indicator in focus influences the safety targets and overall business goals and visions. In Section 7, we refer to the 2005 Texas City refinery accident, and use this to illustrate the main points from the previous discussions. A main reason for referring to this specific accident, is the importance it illustrated for having quality performance indicators for process safety in the refinery and petrochemical industries, for example the developments of API 754 (API Recommended practise 754: 2010; 2016). A performance indicators program provides useful information for driving improvement and when acted upon, contributes to reducing risks of major hazards by identifying the underlying causes and taking action to prevent recurrence. In Section 8, the SPI is assessed by referring to another incident, the 2005 Buncefield oil storage depot accident, to illustrate its usefulness in a broader context. Finally, in Section 9, we give some conclusions, including recommendations regarding the appropriateness of using the SMART criteria in the context considered.

## 2. Measurement of performance in safety management

SPIs are used to provide insights into safety performance, something that is conceptually difficult to measure directly. The indicators are measures that express the level of safety performance achieved for a given system, particularly barriers, and representing a type of key performance indicator allowing for measurable results linked to both quantitative and qualitative findings (ISO 41011:2017). A safety indicator covers any indicator giving relevant information about the state of equipment, organization or human activity related to safety, for example the number of hydrocarbon leakages, which are type of events linked to higher risk for major accidents (Vinnem 2012). Another key indicator measuring barrier safety performance, is the ‘failure fraction’, which is e.g. used by the Petroleum Safety Authority Norway in their analysis of the risk level on the Norwegian Continental Shelf. It gives the ratio between number of failures and the corresponding number of tests performed (Selvik and Abrahamsen 2015). In general, the information achieved through the indicators should be able to help identify whether barrier- or safety-related actions are needed. As such, the use of such indicators are in line with the suggestions from particularly Saleh et al. (2014a; 2014b), pointing to the importance of the “safety-diagnosability principle”, where focus is on the ability to identify dangerous states in the operations through observability. A key is to achieve reliable information about the barrier safety performance, where the selected indicator is suitable for the application and can be used for a meaningful evaluation of the performance.

Barrier management is a core part of the safety management, which in the process industries is about establishing and maintaining layers of protection against hazardous events to achieve specified safety objectives, as part of overall safety management. According to the Norwegian Petroleum Safety Authority, the purpose is “to establish and maintain barriers so that the risk faced at any given time can be handled by preventing an undesirable incident from occurring or by limiting the consequences should such an incident occur” (PSA 2013). It concerns having barriers, i.e. “functional grouping of safeguards or controls selected to prevent major accident or limit the consequences” (ISO 17776:2016), which could be of either technological, organizational or human character. For the technological barriers, terms such as ‘hardware’, ‘process’, ‘process safety’ or ‘process-related’ are often used to label the type of barrier. For the different types there exist also several sub-categorisations. Refer to e.g. NORSOK D-010 (2013), for operations

on the Norwegian Continental Shelf, giving guidance for barriers in drilling and well systems; and, reports from the International Association of Oil and Gas Producers (IOGP 2016; IOGP 2018a), giving general categorisations and description of the “hardware” and “human” barrier types.

The barrier management and use of SPI, are similar to general use of key performance indicators, where the information acquired allows for informed decisions by evaluating the level of past, current or future performance. To support barrier management, multiple indicators (an indicator portfolio) are tracked, as the performance cannot usually be described from only one indicator. For example, regarding the quality of a barrier element, both reliability and maintenance information could be relevant and are normally evaluated. A list of relevant indicators from the reliability and maintenance field are given in Annex E of ISO 14224 (2016), which includes common measures such as the ‘mean time to failure’ (MTTR), the ‘mean overall repair time’ (MRT), and also ‘technical availability’ and ‘operational availability’; see also EN 15341 (2019) guiding the use of maintenance indicators. Such measures are widely used across process industries and the combining of different SPI are important for the overall monitoring of barrier performance and safety management, but also for general business management though the link to safety objectives or goals.

OECD (2008) separates between ‘activities’ and ‘outcome’ indicators, in the context of chemical process barriers. Activities indicators are proactive, meaning that they provide information about ongoing activities and conditions, and/or development of these, expressing the potential of barrier failure or accidents. This type is often called ‘leading’ as the information is supposed to help predicting or giving some expectation about future safety, before anything critical occurs. It is giving answers to ‘why’ safety performance is going in some direction. Outcome indicators, on the other side, are reactive. These intended to provide information about the effects of operations and actions taken, having then instead focus on observable events occurring. It addresses the current or past performance, thus giving answers to ‘what occurred’. Often this latter type is labelled as ‘lagging indicators’; see Kongsvik et al. (2011), Payne et al. (2009), Tamim et al. (2017), Smith and Moblely (2008) and IOGP (2018b). There is also a type called ‘diagnostic indicators’, used for performance indicators that are signal the health of processes or activities (Badawy et al., 2016; Peng et al., 2007). These are not directly linked to potential for safety events occurring, but rather focusing on the general safety culture level.

API Recommended Practise 754 (2016), strongly motivated by the 2005 Texas refinery accident (see Section 7), focus on both activities and outcome indicators. And both types should follow the same basic principles for quality:

- Indicators should drive process safety performance improvement and learning
- Indicators should be relatively easy to implement and easily understood by all stakeholders (e.g. workers and the public)
- Indicators should be statistically valid at one or more of the following levels: industry, company, and site
- Indicators should be appropriate for industry, company or site level benchmarking

It is clearly relevant to capture both activities and outcome indicators when evaluating safety performance. Further, as there are multiple indicators providing input, some structured approach, for dealing with them and combining the information, is required, for example, using balanced scorecards (Kaplan and Norton 1996; Vukomanovic and Radujkovic 2013). The scorecards allow for easier overview of the aspects measured and what tolerance levels the measures are tested against. The evaluation depends on what is the motivation of the indicator(s), beyond having a safety relation. There could be motivations such as:

- Evaluating the ability to meet objectives and safety targets
- Identifying focus and improvement areas
- Monitoring quantitative effect of actions taken
- Demonstrating that some benchmark level is satisfied

The SPIs provide key safety information, which gives them a role also in overall business management. A main task is to establish a link between the information achieved through the set of indicators selected, covering then a portfolio of SPI, and their ability to create overall value and quality in decision-making, where the quality of the SPIs obviously plays an important role.

### 3. SMART criteria overview

The SMART criteria have a broad application area, and are used for various key performance indicators, not only safety or barrier indicators. The reference to these criteria in relation to assessment of quality allows for a transparent process, where each of the criteria needs to be assessed and satisfied. It is a common way of considering quality aspects of information potentially having business value. This because the information links to decisions that influence achievement of goal, targets and visions (Parida and Kumar 2006; Kaganski and Toomalu 2017). By satisfying all five of the SMART criteria, the information provided by the indicator demonstrates usefulness as well as adequate quality. See also Doran (1981), which is often cited in relation to quality, goals and business objectives. For the history of the development of ‘SMART’, we refer to for example Lawlor and Hornyak (2012).

The five criteria are listed in Table 1 along with a brief description on what is covered. When all five criteria are satisfied, then the SPI in principle is having adequate quality to inform decision-making in barrier management.

Above SMART is presented as being “one” specific set of criteria. But in fact, there are different versions of the SMART acronym being used, where the letters could refer to other aspects or criteria. As one example, the letter ‘S’ sometimes refers to ‘sustainable’, ‘A’ sometimes refers to ‘attainable’, the ‘R’ to ‘realistic’ and the ‘T’ to ‘traceable’, but typically, the combination of alternatives suggested in literature covers more or less similar meaning. The following overview gives examples of possible alternatives for the letters used in and applicable for the SMART acronym:

- S: Short; Sensible; Simple; Significant; Strategic; Stretching; Sustainable  
 M: Maintainable; Manageable; Meaningful; Motivating  
 A: Acceptable; Adjustable; Adaptable; Action-oriented; Agreeable; Aligned; Appropriate; Attainable  
 R: Relative; Results-oriented; Rewarding; Reviewable; Robust  
 T: Trackable; Traceable; Tangible; Time- (bound; constrained; constricted; related; specific)

There are also acronym variations, such as for example ‘SMARTER’, which extends with two additional letters and criteria. This is suggested by several, for example Vukomanovic and Radujkovic (2013), Kaufman et al. (2003) and Galligan et al. (2000). The common meaning of the new letters ‘E’ and ‘R’ then being (Better Regulation Task Force 2000) ‘explainability’, meaning that the indicator is simple to understand and communicate; and ‘relativity’, meaning that the indicator is still considered as useful or applicable if business conditions change (for example if production volume increases), respectively. Regarding the new ‘E’, ‘explainability’, it might be argued that this is similar to the criteria ‘specificity’ used for the letter ‘S’. And, the new ‘R’, ‘relativity’, is to some extent is already covered by the criteria ‘relevancy’ being already used for the letter ‘R’, which expresses relevancy in changed business conditions. Hence, there are reasons to question whether the added letters add much or whether these additions are more motivated by a motivation to come up with something new or design a catchy

acronym.

There are also other acronyms that extends 'SMART' by adding just one letter, such as for example 'SMAART', where the letter 'A's could refer to 'attainable' and 'action-oriented'. There is also 'C-SMART', attained by adding the letter 'C' for 'challenging' or 'controllable'. In addition to 'SMARTER', there are also other two-letter suggestions such as for example 'SMARTIE', adding 'I' and 'E', for 'inspiring' and 'enthusiasm'. It is another example of an acronym created to achieve a nice acronym, where the letter 'M' for 'motivating' could have been used instead but would not produce such a catchy acronym. Then, we have the double-layer 'SMART' variants, the 'SMART2' and 'SMART<sup>2</sup>', meaning that each of the letters in the acronym is considered twice (RapidBI 2016; Kavanagh 2013).

For the discussions in the following sections, we will focus the criteria listed and described in Table 1. However, several of the other criteria mentioned above as potential candidates and possibly relevant quality aspects will to some extent be part of the discussion on whether a sufficiently broad quality picture is achieved by using the SMART criteria.

#### 4. Use of the SMART criteria to assess quality of a SPI

In this section we address the assessment of SPI quality when disregarding the portfolio influence. We do not yet assess the influence from other indicators in the portfolio, and we only assess the quality of an individual and isolated SPI. It also means that we are not considering the managerial context and influence in the assessments, and thus fail to consider the broader picture. This simplifies the quality assessments, as there is then no need to cover the portfolio management and possible duplicity or conflict of interest between the indicators included. We leave to Section 5 the discussion related to the quality influence from the way the indicator portfolio is composed. The role of the SPI from a portfolio perspective is obviously important and relevant, but is for now ignored, meaning that a quality SPI, individually, does not depend on how it is used and balanced with other indicators. Hence, we have the situation where a SPI could be acceptable, while the portfolio of SPIs, of which it is part of, could have low quality.

The value of the information provided by the indicator needs to be seen in relation to the decision-making where it applied. However, at the time when the indicator is selected, it might not be clear exactly how it will be used. Understanding how it will be used, makes it possible to consider the value it might have in barrier management. It is about usefulness. A main characteristic of quality in relation to quality decision-making is that the information is useful. According to Matheson and Matheson (1998), as one out of six dimensions characterising quality decision-making:

- Helpful frame (what is it that I am deciding?)
- Creative alternatives (what are my choices?)
- Useful information (what do I know?)
- Clear values (what consequences do I care about?)
- Sound reasoning (am I thinking straight about this?)
- Commitment to follow through (will I really take action?)

Being useful is about applicability for its area of use, but also means that it should be compatible with the data handling tools being used, which is becoming important when dealing with software products, big data, etc. The combination of information provided by the SPI and applicability influences decision quality, and then also influences how such data can create business value. Further, Bratvold and Begg (2010) state that the two aspects 'reliable' and 'relevant' are part of the 'information usefulness'. 'Reliable' referring to both the source, how it is collected, and the content of the information provided. For the information to be 'reliable', it should be unbiased, representative and verifiable, such that the numbers give a correct representation of the situation. These aspects are to some extent already covered by the 'achievability' and 'relevancy'

components in SMART, as then appropriate information is provided, the SPI is of interest to the context considered, and it has the ability or characteristics to influence the associated barrier management decision-making. What it means in practice, is that any indicator that is ambiguous, complicated, difficult to analyse, vague, analyst-dependent, or not linked to business objectives is obviously characterised as of poor quality, and thus not very useful or valuable.

The main question, then, is whether usefulness is adequately covered by the SMART criteria. If not, there is a strong argument for claiming that the criteria cannot be used to demonstrate acceptable quality. We will go through the five criteria and discuss this below.

We start with 'S' for 'specificity'. For the information to be useful, it is difficult to argue against the claim that it should be understandable and clearly expressed. There should not be any room for misinterpreting the meaning or definition of the indicator but be clear what kind of information it provides such that it is interpreted consistently. This relates also to the 'time-based' (T) criteria. There is no point in measuring the performance if the period considered is off. Overall, it is a matter of having precise knowledge. Implicitly then, aspects such as 'consistency', 'explainability', and 'transparency' are also covered.

Moving to 'M' for 'measurability', it could be questioned whether there is substantial need for this criterion, as any SPI, being a measure by definition, is measurable per se. Basically, the safety aspect addressed must be "qualified" as a safety-related measure, we do not see there is a need to include this criterion. Also, the criterion relates somewhat to whether the information needed to perform the calculations are possible to collect or produce with quality, but this is already covered by the following criterion, 'achievability' (A). This one is assumed very important, as it should be possible to produce the numbers with acceptable quality, which is evaluated from this criterion. For example, the calculation should not be overly complex. It could perhaps be better to use the term 'producibility', where it not for it starting with the "wrong" letter and would not give such a catchy acronym. Nevertheless, this criterion opens a way for capturing uncertainty. When including it, it comprises some evaluation of uncertainty regarding the numbers produced.

Finally, we have the 'relevancy' (R) criterion, on whether the indicator information matters to the management of safety performance. It would obviously be possible that it adds value beyond safety, for example provides general business value, but it would then not contribute to the barrier or safety management, which here is the focus and objective. There is a need to state whether the measurement reflects safety or barrier performance, not only measure some changing conditions, i.e. measure according to intention.

For the SPI to have safety-value, it should also be considered so-called 'safety-sensitive', which relates to the 'relevancy' criterion. One could maybe discuss exactly how sensitive the indicator needs to be; however, we find it here sufficient that there is such a relationship, and will not pursue further discussion about the strength here. While we not yet will consider the situation from a portfolio perspective, 'relevant' also means that the particular safety-aspect measured is not already covered by other indicators used. Although there could for some situations be reasonable to include information from two or more indicators on similar aspects, it does not add much value except confirming the results or observations to be correct. Also, it is challenging to conclude on the usefulness of the SPI without considering the other SPIs used. Relevancy is to a large degree a managerial review activity, which cannot be disregarded when evaluating the usefulness of the SPI. For example, it depends on which safety or decision-making principles are adopted and how these are used. This activity involves assessing the whole portfolio, although it is clearly possible to make some decisions based on results from individual SPIs. But, the particular role of the SPI within the portfolio is an issue that is then not covered by the 'relevancy' criterion. Regarding the alternatives for the letter 'R', as indicated in Section 3, some suggest a 'relativity' criterion. We assume that this

criterion is already covered by ‘specificity’ as the situation for which the SPI applies should be precisely described.

To summarize the discussion above (see Table 2), we conclude that all criteria deal with relevant quality aspects. The letters ‘S’ and ‘M’, and some degree ‘T’, refer to ‘what we know’ aspects, the letter ‘A’ refers to ‘how to use it’ aspects, the letter ‘R’ focus on ‘why’ aspects, and the letter ‘T’ refers to aspects related to ‘when or which period to consider’. For ‘measurability’ it may be questioned whether perhaps this criterion could be removed being implicitly already covered, as any SPI by definition qualifies as a measure. In principle, there is not a problem keeping it, but it adds limited value. By including it we just achieve an assessment of whether the safety or barrier phenomena considered, is possible to measure, which is basically the same being assessed by the achievability criterion. However, we await the discussion from a portfolio perspective before making any conclusions on this issue.

The five criteria discussed above seem all relevant to some degree, but there also other candidates that could be considered, to complement the aspects already covered. Neither of the letters links specifically to the aspect of ‘how to use it’, although, it is part of the ‘relevancy’ aspect, as it measures safety or barrier performance and implicitly assumes that a safety or a barrier action is required if performance is for example poor. But it is not fully covered by this. Say, for example, that we consider ‘extreme weather events’ as a basis for a SPI. Would such a measure satisfy all five criteria discussed above? For overall business performance, it might be the situation. But not for safety performance. Clearly, it would not be very useful as a SPI. Yes, extreme weather may have a safety impact, but it will be possible to take precautionary or consequence-reducing measures. For safety and barrier management, any SPI that are checked as ‘relevant’, are implicitly associated with a possibility to make decisions influencing or controlling future outcomes recorded by the measure. Selvik et al. (2020) claim that one key quality characteristics, related to a discussion on key performance indicator quality in general, is that they are controllable. In a safety context, it means that appropriate safety-related actions might have an effect and could improve SPI results, but as that is assumed to always be the situation, we cannot see a need for this criterion. In a safety or barrier context, if we are not able to improve safety with respect to the aspect considered, the indicator is not ‘relevant’ and of minimal usefulness. Hence, as for M in ‘SMART’, we cannot see that it matters much whether ‘manageable’ or ‘measurable’ is selected in the SMART acronym, both add aspects already covered by the other criteria.

An example of the use of the SMART criteria is given in Section 7, where the criteria are discussed with basis in the 2005 accident that occurred in a petroleum refinery in Texas after critical barrier failures. However, we should also consider the role of the SPI portfolio as part of the overall SPI quality assessment in situations where several indicators are tracked. As already stated, we find it insufficient to consider quality without making assessments on what influence the other portfolio indicators have. This is a main aspect of ‘relevancy’. In the discussion in Section 6, we address how the inclusion of other indicators matters for the SPI usefulness. But first we present and discuss fundamentally how

**Table 2**  
Overview of which of the SMART criteria are covered by ‘usefulness’.

Criterion	Covered by usefulness?	Comment
Specificity	Yes	Should be in place for clear understanding of the SPI, and for consistent use
Measurability	Yes, but not needed	It should be possible to compare the SPI numbers scientifically. However, this criterion is already covered as the indicator necessarily is a measure.
Achievability	Yes	The SPI must be producible in a consistent way
Relevancy	Yes	The aspect covered by the SPI should matter to safety or barrier management
Time-based	Yes	The SPI is of limited value if the time aspect is poorly covered

to develop the SPI portfolio.

## 5. How to build an indicator portfolio with adequate quality

The management of SPIs involved understanding the results collected from the individual indicators. This requires some structured way that allows the decision-makers to achieve appropriate balance of the indicators included. The use of balanced scorecards is one way. When establishing this structure, again, focus should be away from the distinction and variety of aspects (spread) covered by the indicators, and rather on, as Oien et al. (2011) also argue, in a safety context, how to achieve a useful collection or portfolio of indicators. We refer also to the discussion on the use of leading safety indicators in Leveson (2015).

### 5.1. Identifying candidates for the SPI portfolio

The starting point for selecting SPIs, is to clarify the safety targets and objectives beyond the barrier requirements. The targets and objectives should be framed for the relevant context, such that the appropriate level of detail and information support needs for decision-making is reflected. The aim is to achieve a set of SPI that can express a broad spectrum of performance, for management to make safety-informed decisions. The SPI candidates are typically referring to failure information, and many are linked to barrier reliability and maintenance area. Such information is typically business sensitive in general, as having barrier failures can have a significant effect on business value. Hence, the indicators are sometimes labelled as key performance indicators or safety key performance indicators, as in e.g. Bellamy and Sol (2012). Several of these are described in the ISO standard on reliability and maintenance data collection and exchange, ISO 14224 (2016), which recommends that the key performance indicators are aligned to the organisation’s objectives for the facility (or operations), and that improvements are identified and implemented in order to achieve the organisation’s planned objectives. It is then appropriate that the indicator portfolio reflect targets and objectives at different levels, such that they cover various levels of the organisation when aligned with other performance indicator selected for different groups of equipment, systems or personnel. This is not an activity driven by the analyst or decision-maker but rather a coordinated activity of stakeholders, including managers and discipline experts, whose opinions all in some way should be captured in the assessment of the alternative measures and their effects and importance.

The task of selecting amongst SPI candidates involves a structured prioritization of which are the important performance aspects. When focus is on barrier performance, there is usually not many failure events occurring. Hence, hence it is clearly fruitful to map also other candidate types. The candidates normally cover a range of both leading (activities) and lagging (outcome) indicators, and diagnostic indicators. The above-mentioned ISO 14224 (2016) provides a list of 34 key performance indicators which are applicable within the reliability and maintenance area. Bellamy and Sol (2012) present an extensive review on SPIs related to barrier management, and in the review go through relevant candidates. Beyond the typical candidates, where in addition, companies also develop specific candidates suited to their needs. It is a quite complex landscape. However, a key is to identify how the safety or barrier performance may be expressed and to link it to the use of the information. There is overall a large amount of literature discussing the appropriateness of performance indicators, particularly the leading ones (Badawy et al., 2016; Swuste et al., 2016). It illustrates how challenging it can be to select amongst the leading indicator candidates. See also discussion in Bellamy and Sol (2012).

A characteristic of the SPIs is the explicit link to safety performance. Many would perhaps characterise them as ‘appealing’ due to the understandable, simple, and compressed way key safety information is communicated. The SPIs comprise key safety information. Hence, it is not surprising that there is a strong link to the use of risk acceptance

criteria (RAC; see e.g. [Hokstad et al. \(2004\)](#) and [Aven and Vinnem \(2005\)](#)). These may also be labelled as safety acceptance criteria, but risk being the broader umbrella. The RACs indicate some aspect of performance related to risk. The different measures used in the process industries for comparison against some RAC can then be considered as a larger set compared with the safety acceptance criteria, which for example does not cover possible cost consequences. Nevertheless, the use and definition of these criteria as part of the objectives and safety targets, is often found as the basis for the selection of appropriate SPIs. For example, an indicator may be selected to assess and evaluate against some defined acceptable criteria.

Focus when addressing the quality of a specific indicator part of a SPI portfolio, is on its value. Without adding value, the information has minimal contribution or is misleading in decision-making and is obviously not considered very useful. For example, SPIs measuring 'wrong things', such as indicators with no 'path' to credible accident events, or is having significant uncertainty, should be avoided. The consideration is closely linked with traditional value of information assessment ([Bratvold et al., 2007](#); [Bjørnsen et al., 2019](#)), analysing and evaluating to what degree the information (here the indicator information as part of the SPI portfolio), has a significant influence on the decision-making. In practise, this is achieved by the indicator having a safety role not already covered by other indicators in the SPI portfolio, for example, by identifying safety or barrier status and trends, and calling for actions. It can be claimed that the indicators should be 'action-guiding'.

As mentioned, selecting amongst SPI candidates involves a structured prioritization. One alternative, which may be used as basis for ranking the candidates for evaluation of which is the more useful, is the use of a multi-criteria analysis. An example is the traditional 'analytical hierarchy process' ([Saaty 1980](#)). Such an analysis is presented by [Elhuni and Ahmad \(2017\)](#) and used to assess 14 different key performance indicators considered for an oil and gas company in Libya. Such a prioritization can be fruitful to identify whether there are candidates with low value. However, despite there are several challenges associated with having large SPI portfolios, as discussed in [Parida and Chattopadhyay \(2007\)](#), there could be good reasons for including many indicators. For example, the operations having many safety facets. In principle, there are no restrictions regarding how many SPIs should be included, as long as the contribution is good. Companies should select the set of SPI candidates that are best suited to their safety objectives and targets. The main principle is that the SPIs combined are contributing with useful information. Obviously then, companies need flexibility as there is not a one solution that fits all. For example, there could be different designs making equipment failures more or less severe, making a big difference for management of the barrier elements across the companies. Target and objectives may be different, as well as digital tools for handling the SPI portfolio; all influencing the portfolio setup. Besides, inside the company there are likely to be sub-organisations with different safety targets and objectives. This giving root to sub-organisations selecting a set of indicators best suited to their needs.

## 5.2. Combining information from the selected SPIs

After identifying individual SPIs with adequate quality, next step is to combine these into an appropriate SPI portfolio, i.e. selecting candidates for a new portfolio or considering candidates to complement an existing one. The challenge is to develop a quality portfolio that is aligned with intended or planned use, as well as targets and objectives. However, this is far from a simple task. A set of SPI candidates are identified, but it is not obvious how to then identify combinations of these giving basis for good decisions, or whether the possible combinations are able to completely cover the safety information needs with respect to the company's safety and barrier management. There is a need to see beyond the individual indicators and understand how they work together, i.e. 'coherence'.

As indicated already, there are different ways of combining the SPIs,

but also different ways to visualise or communicate the portfolio. There has also been some development over time, where digital tools are increasingly important for the portfolio management. The typically tools are digital scorecards, dashboards, and analytic reports.

The digital tools allow for presentation of multiple attributes, where the digitalisation could make it simpler to identify scores for attributes linked specifically to safety. For example, it is possible to add colour coding (e.g. red, yellow and green) to highlight the ones having or should be given higher priority, and also adding information about uncertainty related to the individual attributes. These basically list the scores given for each attribute. But there are also other ways. The information could, as some prefer, into one score, making it easier to conclude based on the results. Another way is to restrict the portfolio to a minimum and low number of SPIs. The challenge is then to select the few ones that can present the key safety information needed. This makes it again difficult to achieve the bigger safety picture and could provide misleading information. To some degree, it depends on the type of business and company considered. But, overall, the practise of having a portfolio with one or only a few SPIs, would not have the simplicity and communicative abilities typically characterising the use of SPIs and key performance indicators in general.

To achieve a SPI portfolio with quality, several aspects should be taken into consideration. Despite having clarified safety target and objectives, and selected indicators according to these, everything is not in place. For example, there is the always reoccurring issue of cost versus benefits. There is usually a cost of acquiring the SPI information, which should be seen in relation the benefits. There is also an issue of uncertainty, i.e. to what extent the information provided is credible. Further, the portfolio should cover a broad spectre of performance aspects but without repeating information for similar aspects. Obviously, key aspects considered as useful to have information about, should be included. However, the challenge is often to make sure that key ones are not missed or which ones to leave out.

Above mentioned the possibility of sub-organisations having different safety targets and objectives. Quite often, this is the situation, where there could be conflicting drivers across the organisation. For example, there could be parts of the organisation focusing on solely on maintenance activities, where safety focus and use of various performance indicators, including general key performance indicators, relates to maintenance activities. These could be contradictory when compared with parts dealing with for example on-site process safety. However, for the company overall, assuming the SPIs being consistent with the business and safety strategy of the company, they could both be appropriate. For example, the indicator 'total maintenance cost' (for a given period) is from the maintenance part's side obviously a number that should be minimised. Seen from an overall company perspective, however, also other aspects that should be part of the consideration. It might be unreasonable to lower the maintenance costs if this leads to significant reduction in reliability and thus higher accident risk. The decision on whether to increase maintenance costs, depends on the reliability and overall safety benefits.

There is an increasing use of digital tools in safety management. There are extensive software applications assisting the analytic tasks and presentation of results. Some of these allow for user friendly interfaces and simplified understandings of safety, however, there is also the challenge that these become sort of 'black boxes' hiding key information, particularly when automated techniques are applied. Nevertheless, such tools allow for also use of machine learning techniques that can be used to identify risk and safety trends (see. e.g. [Bansal et al., 2020](#)), making it possible to identify patterns not else recognisable. Another point is that the use of digital tools makes it possible to reach out and spread information, make it available and useful, in a more effective way. For example, an automated dashboard for SPI tracking could allow for 'real-time' updates. Related to the maintenance activities, such use is associated to 'maintenance excellence' status, meaning that reliability and maintenance performance should be aligned at a strategic level and

the performance should be communicated in an appropriate way. An industrial example is Maersk Oil Qatar's efforts to achieve such status, where the use of effective communication means to present performance aspects is seen as very important (Smart and Blakey 2014). Another example is the 'maintenance excellence' programme built in Shell, for which Jansen (2015) claims that: A "computerized maintenance management system (CMMS) should be the backbone for work management and performance improvement", stating the importance of bridging performance indicators and the digital tools.

Finally, before turning to a discussion on use of the revised SMART criteria, we acknowledge that the safety situation and associated targets and objectives are not a static matter. This is something that could change, for example due to measures implemented or requirement for more robust designs. The indicators should reflect a situation of targets and objectives being dynamically redefined. There is a need to continuously review whether the basis for the SPI construction holds, and if needed, to update the SPI portfolio and reconsider how to use the information, as argued in Øien et al. (2011).

## 6. Using the SMART criteria to assess the indicator quality from a portfolio perspective

Including an assessment of the SPI portfolio complicates the quality assessment. It becomes more complex in nature, partly because the other SPIs might not be sufficiently clear on the spectrum of use (decision-making situations) and usefulness. It is challenging when having to capture a mix of attributes. There is also aspects of confidence and resources needed to perform the quality assessment, not always in place. These are typical challenges, when using the information in safety or barrier decision-making, addressed in the 'managerial review and judgement'. There are likely to be situations where the benefits or usefulness of the SPIs can be questioned, for example because there is not collected a specific type of data or there not being enough history to conclude with certainty. It is not the intention that the SPI should support all types of safety or barrier decisions. The SPIs provide information giving insights into safety or barrier 'performance' and business 'health.' They should not be seen as available 'decision-making instruments.' A fundamental principle of the 'managerial review and judgement' activity is that it is the responsibility of the decision-makers to consider what information is appropriate and how to use this in decision-making situations. It is an activity where management considers and weights the different concerns, including interests from various stakeholders (internal and external). Again, the use of the SPI portfolio is a dynamic process; being strongly influenced by the context and stakeholders involved. As such, quality is interpreted as a relative matter. It is a result of those involved, which obviously could make it challenging to assess the SPI usefulness.

In the same way as for the assessment of individual SPIs (outlined and discussed in Section 4), the assessment should be performed with respect to safety targets, objectives, and usefulness, also when taking a portfolio perspective. Focus is still on achieving or contributing to improved decision quality. However, this requires the safety targets and objectives to be clearly defined. Otherwise it is difficult to evaluate whether the SPIs are useful or needed. Next, we will discuss the use of the modified SMART criteria for the quality assessment.

As in Section 4, we start with the 'specificity' (S) criterion. There is no doubt in this quality aspect being relevant. But focus is slightly different. When considering this aspect from a portfolio perspective, 'specificity' extends beyond the specific SPI in focus and covers also the other SPIs in the portfolio. Hence, for this criterion to be satisfied, there should be precise information on which other SPIs are included, besides, it should be clearly stated how the SPIs are combined in the portfolio and how the information is expressed (pictured). For example, information on SPI ranking or priority should be available, to define clearly the SPI roles in the portfolio and how they compare for decision-making purposes. Such specificity makes it simple to understand the purpose of the

SPI amongst the other SPIs, and how it can be used in barrier and safety management.

Continuing with the next 'SMART' criteria, we have then 'manageability' (M). The point of this criterion is to assess whether, when combined with the full portfolio, there are challenges restricting management of the safety aspects addressed by the SPI in focus. For example, there could be a situation where real safety benefits cannot be achieved as this would 'steal' resources from other and more critical safety activities. In other words, it means that it is in principle manageable, but not in practise. Assessment of the specific SPI as part of a defined portfolio addresses the ability to manage the SPI in focus seen from a systems perspective. The point is not to find a suitable way of managing the portfolio but, rather, to identify what is the room for improvement of the considered safety aspect, given a more relevant context of the current situation. Prioritization of resources and the SPI role could clearly make a difference for this ability. However, this would be a managerial task and for the quality assessment, the conclusion would always be that it is possible to manage safety or barrier performance in some way. As for the conclusion that a relevant SPI is always manageable from an individual SPI perspective, although the actions are not identified specifically, this will also be the situation when taking account also the other indicators part of the portfolio. As the 'M' criterion adds no value to the quality assessment, it would be better, for the safety indicator context, to shorten the acronym to 'STAR'.

The 'achievability' (A) criterion follows up on the managerial (the decision-maker's) ability to take actions. Again, there is a need to consider that the management could be facing several conflicting safety targets and objectives being addressed by different SPIs in the portfolio. Basically, what we need to assess is, whether it is possible to achieve SPI results with adequate quality when combined with the portfolio of SPIs. This implicitly relates to the way the results are integrated in the format used to compile the SPI results, for example using digital scorecards. As for the 'manageability' criterion, the conclusion reached for the 'achievability', is likely to be the same for both the individual SPI quality assessment and for the portfolio SPI assessment. Not necessarily, but usually this will be the situation.

The 'relevancy' (R) criterion is perhaps the one attracting most attention. At least in literature because of the strong link to 'why' the company should spend resources on it. The assessment of this covers the ability to make good safety decision and take appropriate actions using the information from the available multi-attribute indicator portfolio (Wood 2016; Longhi et al., 2015). Quality, then, comes from whether the decision-makers are able to make safety-informed decisions showing a positive effect on the performance aspect considered, which are based on the information provided by the SPI(s), and would not have been made otherwise. From an individual SPI perspective, this criterion is already considered; however, there is again the possibility that changes to the specific indicator, could have an overall negative effect on safety performance when also other SPIs are considered, for instance, a conflict of interest could exist between the SPIs. Hence, we could have a situation where it is possible to manage the SPI results over time, but where the benefits of the specific indicator are marginal or disproportionate compared with the benefits obtained from the portfolio. For example, it could be that the safety aspect in focus is already covered, or partly covered, by another SPI.

Related to information needs in various management situations, there is often assumed a relationship between management and measurement in line with the saying, that: "you cannot manage if you don't measure". It is about having enough information to make good decisions and to have some level of control over the situation. However, related to performance measurements, associated analysis and decision-making, we often find the opposite to be just as relevant: "what you measure is what you manage". The information and knowledge obtained from the SPIs could assist in establishing a safety picture describing the current situation, but clearly this information may also have strong influence on which safety aspects are given priority. Say the company has adopted a

vision zero principle, i.e. defining a safety target and vision of zero critical personnel injuries and fatalities. Then, based on this, SPI could be developed to track the number of events occurring and use this information to guide further improvements. However, management guided from this SPI, despite being suited to this objective, could fail to be rational if it is compared with traditional cost-benefit principles and overall safety benefits, i.e. seen from a system perspective.

'Time-based' (T), being the final criterion, considers whether the defined measurement period is appropriate, when used in combination with the other SPIs. An argument for considering a different period, is that similar information is already provided by another SPI. It could be appropriate to make changes, to make the portfolio cover the complete range of past, present and future performance. In a similar way, the portfolio should cover target and objectives of both operational and strategic character, i.e. short-term and long-term, respectively.

## 7. Use of the modified criteria (STAR) to assess a safety performance indicator in a refinery scenario

In this section, we will consider a safety performance indicator called 'Dangerous fluid overflowing events.' This indicator could be attractive to process industries and is obviously related to safety. Monitoring of trends and level of occurrence can potentially add value by identifying undesired safety and business performance. According to [Chang and Lin \(2006\)](#), overflowing events cause a loss of containment and claim it to be the most frequent cause of operational error for tank accidents. Overflow hazard also depends on the type of vessel and associated upstream/downstream equipment ([Summers and Hearn 2010](#)). There are differences in the fluid overflowing for a process vessel vs. storage tank. The distinction between the two types of equipment is clarified e.g. in [ISO 14224 \(2016\)](#), which details taxonomy classification for reliability data collection within the process industries. Both are listed as a mechanical equipment category and show that storage tanks and pressure vessels contain similar subunits. Further, this international standard clarifies that storage tanks include atmospheric tank and low-pressure tanks, while the pressure vessels could handle gas or other fluids with higher pressure.

When a process vessel starts overflowing, usually the fluid outlet of the vessel (e.g. relief system, control valves, etc.) is blocked during the fluid inflow. In a storage tank, an unchecked rate of inflow accumulates large amount of fluid such that it exceeds the tank's maximum holding capacity. After a processing vessel is overfilled, the excess liquid unintentionally enters the outlets designed for gas phase or is passed to the downstream equipment that is not designed to receive it ([Summers and Hearn 2010](#)). An overfilled storage tank releases excess liquid through its vents or fails under excess structural pressure ([Waite 2013](#)). While overflowing may materialize somewhat differently in both vessel types, the overflowing event equally threatens the operations' safety in both. We will investigate the SPI's usefulness in tracking both of these two different conditions.

A main example of an overfilled process vessel is the major accident that occurred at a refinery in Texas City March 2005, where 15 people were killed, 180 was injured, in addition to major structural and financial consequences, from fires and explosions caused by overflowing.

We will use the Texas City refinery scenario, and more specifically the 'Isomerization unit' (ISOM), which was the source of the accident, as basis for the discussion regarding the quality of the overflowing indicator for process vessel. The refinery had previously ignored a past trend of minor-overflowing events assuming it not to pose any hazard, but by that repeatedly removing a key safety barrier. This allows for a discussion on the indicator usefulness from a realistic safety management view, both from individual and from a portfolio perspective. This accident is particularly relevant to assess if the information conveyed by the chosen indicator can help in determining why the combination of safety barriers did not function properly. But before discussing this, we will give a brief and simplified description of the system and what happened. For a more

detailed description, we refer to e.g. [Saleh et al. \(2014b\)](#), [Hopkins \(2008\)](#) and [CSB \(2007\)](#).

### 7.1. Key barriers related to operation of the Texas City refinery ISOM - and what went wrong

[Fig. 1](#) shows a simplified layout of the main components of the ISOM unit at the refinery. Liquid raffinate flows into a tank or vessel called the 'raffinate splitter tower', being the centre of the unit. The vessel is a about 50 m high and is where heavier raffinate is separated, sending parts of the raffinate to storage. The tower has sight glass and a level transmitter (sensor) measuring the fluid level in the range 1.5–2.7 m above bottom. In addition two separate level alarms are installed to indicate high liquid level. The first alarm is programmed to sound when the transmitter's reading reached 2.3 m in the tower. The second alarm is a redundant high-level switch that sounds at 2.4 m fluid level, independent of the level transmitter. The 'level alarm low' is another low-level redundant alarm. From the top of the tower, lighter raffinate flows out and into an air-cooled condenser, from where it is sent either for storage or routed back to the tower. To effectively deal with potential high level or over-pressurisation, upset operations or shutdowns, three parallel safety relief valves are installed. The outlet of this line leads to the disposal system, i.e. 'blowdown drum and stack' and 'sewer'. Liquids will then end up at the bottom while, the gases escapes to air through the vent stack on the top. The liquids then discharge into the unit's sewer by opening a manual block valve. The blowdown drum had level sight glass for level monitoring and a high-level alarm to alert operators when liquid was close to flowing above a certain level (i.e. seal leg of the gooseneck pipe opening to the drain).

On the morning of the accident, when starting up, the lead operator as usual started pumping raffinate into the splitter tower. According to plant operators' common practise, although a violation of formal start-up procedure that calls to maintain 50 percent transmitter reading level, the raffinate was pumped in to a 99 percent transmitter level. As the tower was filling up beyond the set point of the high-level alarms, only one high-level alarm triggered but was ignored. The redundant high alarm did not sound. The level sight glass was not readable and not used. The operator was unaware and interpreted the transmitter's 99% (maximum) reading as the correct level measurement. In reality, the tower had filled 1.2 m above the top level of the transmitter's range. After the raffinate section equipment were filled up, the start-up procedure and raffinate feed were suspended. Against procedure, the operator also closed a control valve instead of leaving this in 'automatic' mode. Before, leaving, the night shift operator left incomplete information in the logbook about what steps were taken and what was to be done in the next shift.

Consequently, the next shiftoperator did not receive proper information about the unit's status. Due to the miscommunication, the new operator was unaware that the raffinate equipment was filled during the previous shift. The unit supervisors were also unaware of these conditions. Next morning, due to miscommunication, the supervisors instructed the operations crew to restart the raffinate feed into the tower. The operators controlling the heavy and light raffinate products were uncoordinated. They did not receive clear instructions about the feed and product routing prior to start-up. They made false assumptions about the conditions and ended up closing both the level control valves (outlets) while the tower was continuously being fed. The splitter tower was unknowingly being overfilled now as it had no output discharge or real level monitoring. At the time when the operator raised the temperature of raffinate in the splitter tower, the level transmitter falsely displayed 2.6 m fluid level (investigation reports indicate the level was in fact around 20 m and increasing). Some hours later, the overflowing was still unknown to the operators, who still misinterpreted the system behaviour. It ultimately led to raffinate liquid overflowing to the overhead line, through the safety relief valves and into the blowdown drum. And, without the operators knowing it, the blowdown drum filled up



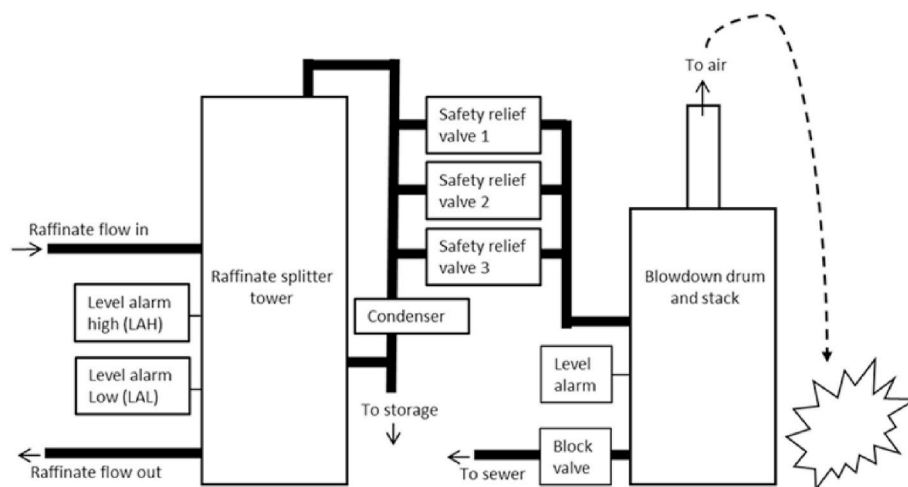


Fig. 1. ISOM unit – Simplified layout.

(the level alarm was out) and raffinate was shot out through the vent stack into the air. At the ground, vapor ignited, most likely from a nearby idling pickup truck, causing a massive explosion. Clearly, a series of safety barriers for preventing dangerous fluid overfilling failed on the way; see below.

#### 7.1.1. Organisational safety barriers

Operators and staff controlling the ISOM unit, was inadequate. They were overworked and poorly trained to handle the abnormal start-up conditions leading to fluid overfilling. The control room was ill-equipped to display the net fluid flow rate or to detect overfilling events. There were insufficient instructions to the operators regarding how to consider the incoming-outgoing raffinate flow readings being essential for overfilling situations, and particularly relevant during start-ups. The company to large extent failed in enforcing formal procedure (e.g. inadequate shift handover, poor recording quality in logbooks, lack of technical supervision, no instrumentation checks pre-start-up). There was also a history of budget restrictions delaying maintenance activity. Overall, the organisational barriers of promoting a safety culture, providing adequate safety preparedness and operator training were largely failing.

#### 7.1.2. Human safety barriers

The operators frequently ignored alarms at the unit and violated start-up procedures. Besides, there was a lack of communication among the shift operators and management in conveying critical decisions, such as the decision not to follow formal start-up procedure. The human barriers of skill, training and experience failed to detect the overfilling incident and containing it early.

#### 7.1.3. Technical safety barriers

The instruments were poorly calibrated or not designed to detect the actual fluid level. The sight glass needed replacement, and the high-level alarms failed to activate, both at the tower and at the blowdown drum. The failure of the level alarms meant that the operator received no warning about the critical fluid level nor that it was exceeding detectable level. The sight glasses were both able to only display fluid level in a small range and was poorly designed. The tower's level transmitter was unreliable (e.g. it wrongly displayed fluid below 100% level (2.4 m) when the fluid was overfilling in the tower). Since, the operators trusted this instrument's reliability, they could not detect that the fluid had surpassed the transmitter's recognisable range and was escalating into an overfilling event. The ISOM unit discharged the flammable raffinate into a sewer, however, as per the industry guidelines this was an unsafe practice to prevent blowdown drum overfilling. The system lacked

screening points of fluid flow in and out of the equipment. These weak barriers of instrumentations and alarm systems in combination failed to detect the overfilling incident, making the overfilling go undetected, up to the explosion.

#### 7.2. Quality assessment of the safety performance indicator: dangerous fluid level events

The event described above represents only one event. What we are questioning is, whether it is useful to record the number of such events as a key indicator of safety performance. Below, we will assess the quality of the dangerous fluid overfilling event indicator using the modified SMART criteria, now referred to as the 'STAR criteria'. We will do this both individually and at a portfolio level. For the portfolio level, we adopt relevant SPIs suggested by the CSB accident investigation report (CSB 2007). Note that the adopted list of indicators is selected for the purpose of the discussion in this article, is not meant to be neither exhaustive nor fully representative of any real portfolio of SPIs tracked by the current facility management. There are obviously other relevant candidates not included. The portfolio consists of the following six indicators:

1. Personal fatality and injury rate
2. Days away from work
3. Hazardous material release events
4. Dangerous fluid overfilling events
5. Raffinate pressure indicator
6. Raffinate level indicator

We maintain that this portfolio is dedicated to managing the overall safety performance at the ISOM unit of the refinery. The aim is to use information from these SPIs to manage safety performance and avoid accidents in the future. The discussion regarding the usefulness of the indicator, i.e. 'dangerous fluid level events', is given within this frame.

##### 7.2.1. S - Specificity

To satisfy for 'specificity' the indicator should be defined appropriately. In process industries, vessel 'overflow' is given a comprehensible and specific definition in API 2350 (2012), as the point when the product inside a tank rises to the critical high level i.e. the highest level in the tank that the product can reach without detrimental impact, e.g. product overflow or tank damage (Roos and Myers 2015). The important term being 'critical high level'. The API 2350 (2012) calls this the 'overflow level', which is the maximum fill-level of a product within a tank measured from the gauging reference point, above which level any

additional product will overflow and spill out of the tank. Staying consistent with the standard, all combustible and flammable liquids are under focus because their mismanagement poses a higher safety risk. We refer to these as 'dangerous fluids' or simply 'fluids' in this context.

An overflowing event is thus an event where some vessel is filled with a fluid quantity that is more than the maximum capacity. All situations where the vessel is over-filled, or the operator losing fluid level control to cause spillage or tank damage, should be recorded for the indicator in focus. This allows for making trends over fixed intervals, e.g. annually.

Considering the other SPIs in the portfolio, none of these conflict with the dangerous fluid level indicator. These specifically addresses other safety aspects. Indicators 1 and 2 are mainly concerned with on-site personnel. They are type of indicators tracking occupational safety and standards of the working environment. They provide limited information relevant for process safety. Indicators 3 and 4 are both lagging indicators, recording past safety performance. As material release is not seen as relevant to the ISOM unit, the two should not be overlapping. Indicators 5 and 6 are leading SPI related to process health in the splitter tower, managed in real-time. These two reflect the current system state (pressure and fluid level) and are used by operators to make short-term control decisions. From a portfolio perspective, the SPIs are sufficiently specific on which of the SPIs that are to be prioritized for short-term vs. long-term decisions and to be used to track business and safety goal achievement. We conclude that the 'S' criterion is sufficiently satisfied from an individual and portfolio perspective.

#### 7.2.2. T - Time-based

The indicators should show trend for reasonable timeframe. Hale (2009) claims this motivates appropriate safety actions. The overflow indicator counts the events occurring during the period. The question is whether, for the period considered, there are enough events to produce a meaningful rate (Hopkins 2009). If this period is too short, a lack of events could be mistaken for a sound barrier performance. On the other hand, if long time goes by without any event being recorded, Hopkins (2009) argues that it is not possible to compute a meaningful annual rate, nor is it possible to conclude from one occurrence that safety is deteriorating. The time interval considered should be sufficiently long to capture the system's safety status before and after a safety barrier is deployed so that performance comparison is meaningful. According to API Recommended Practice 754 (2016), recommends reporting indicators by current year process-safety-event count, and a 5-year rolling average on a company and industry level. A 5-year rolling average may perhaps capture a broader spectre of events. Although, by producing the overflowing events with an annual rate, it should be easier to identify outliers, and it should be sufficient to capture a trend.

The SPI portfolio covers a combination of short- and long-term focus. Indicators 3 and 4 is to some extent long-term oriented, by considering achievement of objectives through annual (un-averaged 5-year trend can also be relevant) observation periods, while short-term policy goals are more relevant from indicators 1 and 2. The current system state is observed by indicators 5 and 6, although this information could be of interest also for longer terms, and vice versa for the other indicators. The combination of indicators in the portfolio facilitates observing the operational (process safety) objective achievement and the effect of strategic changes in safety and business policies. From a portfolio perspective, the measurement period is quite flexible and can be changed if required. Overall, when the overflowing indicator is recorded for an annual interval, it sufficiently satisfies the 'T' criteria.

#### 7.2.3. A - Achievability

Achievability refers to the ability to produce accurate information. Which can be challenged by uncertainty regarding the number of events recorded. Basically, the number of events come from recording the instances when the level transmitter show 'overflow/high-level' or by other observation or alarm. However, identifying and segregating an overflow-event is not that straightforward. There are several reasons. According to

Summers and Hearn (2010), operators rarely track the fluid levels directly because a 'high-level' event is an overflowing hazard only when the liquid begins flowing to equipment such not designed to receive it. This is when the overflow event can cause loss of containment, as in the Texas City refinery accident. An overflow may occur in a few minutes or may take several hours. As the event propagation time can vary significantly, its classification becomes uncertain, raising data credibility issues. Besides, the cause of a fluid 'high-level' event depends on the operation mode (i.e. start-up, normal or abnormal) as it can influence the amount of fluid accumulated (Summers and Hearn 2010). For example, a higher level under abnormal conditions could be intentional and necessary to prevent equipment stresses. Making it unclear whether the overflowing event is to be recorded if it is assumed as non-hazardous.

The indicator does not separate between hazardous events and inconsequential overflowing events. Although it may be relevant to analysis, information about operating levels, operational modes, safe-fill levels, etc. are ignored when collecting data for the indicator. In the Texas City accident, the operators accepted a high-level against the prescribed start-up procedure. This was due to a lack of information on the safe-fill limit and the level transmitter displaying a limited operating fluid-level range. But assuming the raffinate level in the vessel to be below the high level. Ignoring the role of the measuring device, crucial for this indicator, may produce uncertain and misleading results. A limited-range or unreliable transmitter can result in failure to identify overflowing events in some situations, and perhaps include non-events in others.

The key is to collect credible information about the barrier performance. But as claimed in Saleh et al. (2014a,b), the design configuration and equipment limitations, challenge the ability to collect such information with high credibility. Basically, the uncertainty is significant, making the indicator subject to phenomenon understanding, as it is necessary to assess this uncertainty. A peer-group trend comparison would be risking using misrepresented data. Such an indicator could motivate mistargeted actions, clearly not being in line with the safety objectives. Consequently, on an individual basis, the indicator does not satisfy the 'A' criterion within the current design solution.

From a portfolio perspective, it can be discussed how the overflowing indicator is linked to the collection of data to the 'raffinate level' indicator (Indicator 6). If the quality of any of these are good, then it can be assumed that an overflowing will be detected. However, for the system considered this is not fully the situation. However, this relates also to budget restrictions and priorities, as it is possible to implement a way more credible level monitoring system for the vessel. Which would make the overflowing indicator satisfying the 'A' criterion on a portfolio level.

#### 7.2.4. R - Relevancy

Relevancy is perhaps the most important criterion, indicating why to use the indicator. Fluid overflowing is one of the most commonly occurring instances causing near-misses and loss of containment accidents. In the chemical and petrochemical industries, the loss of containment of a hazardous substance has been the main factor in several major incidents (Collins and Keeley, 2003). It is acknowledged that fluid overflow events pose risk and should be given attention sooner. It is a way of measuring the effectiveness of the control upon which the risk control system relies, which is a key according to Hopkins (2008). A high fraction of historic overflowing events analysed by Chang and Lin (2006) ended with fires and explosions, potentially causing major accidents. Chuka et al. (2016) presents a variety of consequences related to containment loss in the process industries.

Dangerous overflowing events as a lagging indicator can be criticised for not giving early warnings, requiring looking further back in the causal chain, at the underlying causes and the condition of the factors that leads to accidents (Øien et al., 2011). However, Hopkins (2009) argues that in situations when hazardous events are occurring frequently enough to produce a meaningful rate, the rate can be used to measure and manage safety. If the events are rare, it is not that relevant, and we

must look to more frequently occurring precursor events to be able to measure safety (Hopkins, 2009). For the refinery scenario we assume there is a significant number of events. Historic data showed that the processing tower experienced dramatic swings in liquid level during 18 of the 19 previous start-ups and had numerous tower overfilling incidents (CSB, 2007). Between 1995 and 2005, the refinery had four other serious releases from the ISOM unit blowdown drum that were unignited ground-level vapor clouds (The BP US Refineries Independent Safety Review Panel, 2007).

Overfilling events typically follow a complex escalation path, aided by hidden latent failures at different operational stages, which is only implicitly revealed by the overfilling indicator. It does not give the analyst any information about what, where and how the overfilling took place. He must find this out by collecting supporting information (or other SPIs) that underlying conditions and safety gaps. In practise, a variety of safety barriers (e.g. human, technical, organisational) can play a role in preventing such events occurring.

From the portfolio perspective, as Indicator 6 provide a different type of information, i.e. on the current condition, the overfill indicator is complementing the portfolio. Neither the hazardous material release indicator provides conflicting information, as the overfilling refers specifically to the vessel safety performance. This makes the two indicators even more relevant when considered together.

The The BP US Refineries Independent Safety Review Panel (2007) concluded that the operating company in a way placed more attention on personal safety compared with process safety; mistakenly seeing improvement of personal injury rates as an indication of acceptable process safety performance at the refinery. From a portfolio perspective, we can assume that the delay in maintenance actions can be attributed to prioritization of personal safety, promoted by Indicators 1 and 2. It suggests that resources, investments and attention were 'stolen' away from maintaining the overfill-prevention barriers, e.g. installing reliable fluid level transmitters and adequate operator training. The potential for overfilling, on a portfolio level, clearly ranked behind personal safety-targets for the management, as visible in the maintenance budget-cuts, degrading infrastructure, and under-staffed operations (The BP US Refineries Independent Safety Review Panel, 2007). This in practise challenges the benefits at the portfolio level, but also shows why it is important to include such an indicator.

### 7.3. Refinery scenario findings

To summarize the overall results of the above STAR criteria quality assessment of the dangerous-fluid overfilling indicator, we find that there is only one criterion that is not satisfied. The assessment and associated discussion conclude that the criteria specificity, time-based and relevancy are all satisfied. Both for the individual and the portfolio perspectives. However, not the achievability criterion, which fails on both perspectives. Hence, we overall conclude that the indicator in focus is not having adequate quality. This is not to say it cannot be useful, but the achievability obviously challenges this.

## 8. Assessing the safety performance indicator in a storage tank scenario

Above we discuss the overfilling indicator in relation to the Texas City refinery accident, i.e. for a process vessel context. In this section, a similar event i.e. the Buncefield depots' tank overfilling accident is considered. In this accident, the level measurement device did not display the changing level even though the tank's fluid level was rising. This presents a different use case that can be tracked using the indicator. We will re-assess the SPI using the Buncefield case to determine whether it produces similar results on STAR criteria when focusing on storage tanks. It will provide a broader understanding related to the use of this indicator within the process industry.

### 8.1. Key barriers related to operation of the buncefield depot - and what went wrong

Buncefield oil storage & transport depot is a farm of several tanks serving areas in UK, including London. The operating site stored hydrocarbon fuel received via a complex network of three pipelines. It had experienced a devastating explosion and fire in 2005 due to failure of its overfilling protection system.

There were two main safety barriers against tank overfilling. First, an automatic tank gauging system (ATG) that displayed the fuel level on control room screen for the operators to monitor. The ATG also had alarms at 3 succeeding levels (1) 'user high'- set by the supervisor indicating the need for intervention, (2) 'high level'- at a level below the tank's maximum working level, (3) 'high-high level'- at a level above 'high-high' but below IHLS (COMAH 2011). Independent high-level switch (IHLS) was the second barrier set above the ATG alarm levels. Its function was to raise audible alarms when the fuel reached an unintended high level and automatically operated the shut-off valves to stop the fuel supply. IHLS and ATG operated independently of each other to safeguard against tank overfilling. The barriers are illustrated in Fig. 2.

On December 10, 2005, a pipeline started delivering fuel to a storage tank at the depot. But unknowingly the level monitoring instrument of the ATG stopped registering the rising fuel level midway of the delivery. The monitor erroneously displayed a 'flatline' (indicating that the tank was no longer filling up) while the fluid continued to be delivered. The ATG alarms, dependent on this level monitor, could not operate since the level reading remained below their corresponding set levels. The tank's first safety barrier against overfilling had failed. The second barrier, IHLS, was also ineffective because those who installed and operated the switch did not fully understand its working; such that the switch was left effectively inoperable after a previous test (COMAH 2011). The inoperable IHLS meant that neither the final alarm alerted the operators about overfilling nor the automatic fuel supply shutdown activated. Tank's maximum fuel capacity was soon exploited, thereafter the excess fuel started spilling from vents in its roof. This exposed fuel formed a white flammable vapor cloud at the site. After an employee noticed the cloud, he raised an alarm and the firewater pumps got initiated. Almost immediately, the vapor cloud ignited with an explosion of high over-pressure. The explosion was followed by a five-day long fire that injured forty people, engulfed twenty fuel tanks, and had widespread environmental consequences. The overfilling incident was important in causing a complete loss of primary containment (i.e. the tank unit). The failure of ATG recognizing the hazard i.e. misleading level monitors and inoperable IHLS were the main cause for the fuel tank overfilling.

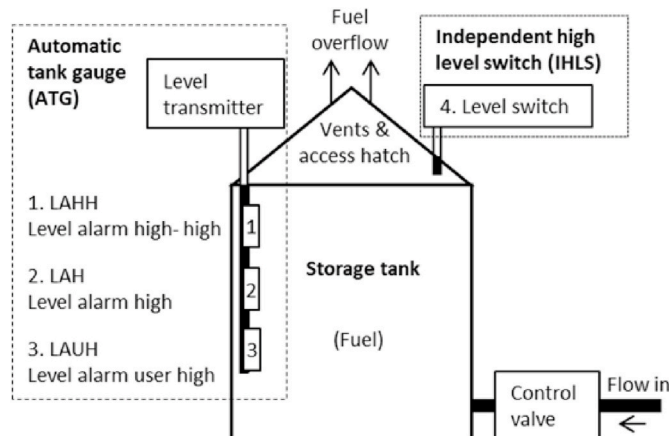


Fig. 2. Buncefield storage tank – Simplified layout.

## 8.2. Quality assessment of safety performance indicator: dangerous fluid level events

The event described above presents a case study of storage tank accident to evaluate the usefulness of recording the overfilling incidents to improve the safety performance. The SPI is already assessed in section 7.2 for its usefulness on the STAR criteria for the case of Texas City refinery's process vessel. Using the results derived from the previous discussion, we reinstate that the overfilling SPI satisfies the 'specificity' and 'time-based' criteria since these qualities are independent of its application.

Next, the 'achievability' of the indicator needs to be examined for storage tanks. As discussed in 7.2, there are uncertainties associated with investigating if a 'high-level' reading indicates an actual overfilling event in the process vessel case. This applies to the storage tanks as well. In the Buncefield case, the tanks had three alarm levels starting from the lowest 'user-level' alarm, raised the need for human intervention incrementally. However, given the poorly specified filling procedures, the Buncefield operators used these alarms subjectively. They underestimated the likelihood of overfilling event by allowing the 'high level' and even 'high-high' level alarms to pass unchecked sometimes (COMAH 2011). The ATG barrier alarms were not being used for performing the intended safety function. The shutdown IHLS barrier was neither properly maintained nor understood clearly. Investigation from the past storage tank accidents commonly point to factors such as poorly maintained hazard measuring devices (alarms and sensors), inconsistently used reporting (logging) system for overfilling incidents, over-worked staff, and lack of data with quality. These factors along with the system complexity and equipment's limitations (refer to section 7.2) add significant uncertainty about the indicator's trend. Therefore, on an individual basis the SPI fails to satisfy the 'A' criterion even for the storage tank application. On a portfolio basis, tank overfilling events can be detected and recorded with the help of other quality indicators.

For the SPI's 'relevance', the consequences of the operation being tracked is important. Fluid filling is the primary operation conducted on a storage tank, often several times every day. Frequent transfer of dangerous fluids warrants monitoring the overfilling events and consequently, its safety barriers' performance. This makes the overfilling SPI particularly relevant for tracking the trend of poorly performed filling operations. Storage tanks are also vulnerable to similar negative consequences of fluid overfilling as discussed for process vessels in 7.2. At the portfolio level, the indicator may receive less or more resource prioritization depending on the management's decision-making principles and risk appetite. In the Buncefield accident, the indicator was ignored by the management and operators alike, as is evident from the investigation report (COMAH 2011). It states that the defect in the tank's level monitor, that had stuck 14 times within three months before the accident, was treated with quick fixes only. The management and staff had underplayed the importance of monitoring key safety trends and later faced the consequences. So, on a standalone as well as portfolio basis, the SPI satisfies the 'R' criterion.

## 8.3. Fuel tank depot findings

To summarize the STAR criteria assessment of dangerous-fluid overfilling indicator, again only the 'A' criterion is unsatisfied for the storage tank application. The indicator is specific, timely and relevant from individual and portfolio perspectives. This case study provides a broader context for SPI's usefulness in a different context. While the indicator's usefulness can be challenged from the aspect of 'achievability', all the other criterion, especially 'relevance', stands in support for the value it can generate for safety barrier's performance management.

## 9. Conclusions

The quality of a safety performance indicator relates to the potential use of this to identify safety challenges for the system considered. This by providing information not already being produced by other indicators, and as such it complements the SPI portfolio. Properly defined and understood indicators can give companies confidence that the right things are being managed and tracked (API Recommended Practise 754:2010).

In this article, we discuss the use of SMART criterion for the quality assessment. This covers five basic criteria assumed to be fruitful for a general key performance indicator context. The SMART criteria cover a range of aspects, which we have considered; one by one. Both individually and from a systems (portfolio) perspective. Overall, we find the criteria to be applicable, and should be included for a general assessment of SPI quality, except for the 'M' aspect. This, regardless of whether the letter 'M' refers to 'measurability' or "manageability". In either of the criterion is assumed to be covered by the other four. We claim that the 'M' can be effectively removed, for both individual and portfolio assessments. Thus, we suggest to instead, when dealing with indicators related to safety business objectives, to rather adopt the following acronym:

'STAR': 'Specificity' – 'Time-based' – 'Achievability' – 'Relevancy'.

The criteria represented by these four letters are suggested as the basis for assessing SPI quality. To demonstrate the use, we have assessed a potential indicator called: Dangerous fluid overfilling events. The assessment identifies significant uncertainty related to producing accurate SPI numbers, and the SPI thus fails for the 'achievability' criterion. The uncertainty, although the indicator is found to be both specific, time-based and highly relevant, challenges the usefulness. Without providing sufficient accuracy it is difficult to use it for informed decision-making and safety business management. However, by using such an indicator there is a chance that one could have seen the 'top of the iceberg' and acted on that to improve the barriers. Besides, as the indicator is seen as highly relevant, this could motivate actions to make it achievable. Overfilling clearly represents a risk, as demonstrated by the 2005 Texas City refinery and the Buncefield depot accident.

The dangerous fluid overfilling indicator assessed is associated with a common safety concern among petroleum but also petrochemical and natural gas industries, as well as nuclear, basically any industry that handles hazardous fluids, i.e. the risk of loss of containment. However, the discussion about quality and usefulness is restricted to the frame and specific system considered and is thus not automatically transferrable to any other process system. Even for other refineries the conclusion could be different. Nevertheless, the use of the STAR criteria is applicable to basically any industry and system being safety oriented.

## Author statement

We confirm that all authors have seen and approved the final version of the revised version of the manuscript.

We also confirm that the article is original work of the authors, hasn't received prior publication and isn't under consideration for publication elsewhere.

## Declaration of competing interest

There are no conflicts of interest to report for the article.

## Acknowledgements

The authors are grateful to the anonymous reviewers for their useful comments and suggestions to the original version of this article.

## References

- I: Safety and Reliability: Methodology and Applications: Proceedings of the European safety and reliability Conference, ESREL 2014, Poland, 14–18 september 2014. Taylor & Francis. ISBN 978-1-138-02681-0. s. 1439–1446. I: Safety and Reliability: Methodology and Applications: Proceedings of the European safety and reliability Conference, ESREL 2014, Poland, 14–18 september 2014. Taylor & Francis. ISBN 978-1-138-02681-0. s. 1439–1446.
- API 2350, 2012. Overfill Protection for Storage Tanks in Petroleum Facilities, fourth ed. Petroleum Institute (API).
- API Recommended Practise 754, 2010. Process Safety Performance Indicators for the Refining and Petrochemical Industries, first ed.
- API Recommended Practise 754, 2016. Process Safety Performance Indicators for the Refining and Petrochemical Industries, second ed.
- Aven, T., Vinnem, J.E., 2005. On the use of risk acceptance criteria in the offshore oil and gas industry. *Reliab. Eng. Syst. Saf.* 90, 15–24.
- Badawy, M., Abd El-Aziz, A.A., Idress, A.M., Hefny, H., Hossam, S., 2016. A survey on exploring key performance indicators. *Future Computing and Informatics Journal* 1 (1–2), 47–52.
- Bansal, S., Saadalla, N., Selvik, J.T., Abrahamsen, E.B., 2020. Development of a bivariate machine-learning approach for decision-support in offshore drilling operations. In: *Proceeding of the 2020 European Safety and Reliability and Conference, Venice, Italy, November 1-6, 2020*.
- Bellamy, L., Sol, V.M., 2012. A Literature Review on Safety Performance Indicators Supporting the Control of Major Hazards<sup>1</sup>. RIVM Report 620089001/2012. National Institute for Public health and the Environment.
- Better Regulation Task Force, 2000. Principles of Good Regulation. Cabinet Office, London.
- Bjørnsen, K., Selvik, J.T., Aven, T., 2019. A semi-quantitative assessment process for improved use of the expected value of information measure in safety management. *Reliab. Eng. Syst. Saf.* 188, 494–502.
- Bratvold, R.B., Begg, S.H., 2010. Making Good Decisions. Society of Petroleum Engineers, Richardson, TX, USA.
- Bratvold, R.B., Bickel, J.E., Lohne, H.P., 2007. Value of Information in the Oil and Gas Industry: Past, Present and Future. In: *SPE Paper 110378, Presented at the SPE Annual Technical Conference, Anaheim, CA, USA*.
- Chang, J., Lin, C.-C., 2006. A study of storage tank accidents. *J. Loss Prev. Process. Ind.* 19, 51–59.
- Chuka, C.E., Freedom, I.H., Anthony, U., 2016. Risk assessment and consequence evaluation of loss of containment in process industries. *International Journal of Modern Studies in Mechanical Engineering* 2 (1), 14–28.
- Collins, A., Keeley, D., 2003. Loss of Containment Incident Analysis. Health & Safety Laboratory.
- COMAH- Control of Major Accident Hazards, 2011. *Buncefield: why did it happen? Health and safety executive UK*. <https://www.hse.gov.uk/comah/buncefield/buncefild-report.pdf>.
- CSB – Chemical Safety and hazard investigation Board, U. S, 2007. Investigation Report, Refinery Explosion and Fire, BP-Texas City, Texas, March 23, 2005. CSB-Report no. 2005-04-I-TX.
- Doran, G.T., 1981. There's a S.M.A.R.T. way to write management's goals and objectives. *Manag. Rev.* 70, 35–36.
- Elhuni, R.M., Ahmad, M.M., 2017. Key Performance Indicators for sustainable production evaluation in oil and gas sector. *Procedia Manufacturing* 11, 718–724.
- EN 15341, 2019. Maintenance - Maintenance Key Performance Indicators. European Standards (EN).
- Galligan, F., Barry, T., Crawford, D., Howe, D., Maskery, C., Ruston, A., Spence, J., 2000. *Advanced PE for Edexcel Student Book*. Heinemann Educational Publishers, Oxford, UK.
- Hale, A., 2009. Editorial special issue on process safety indicators. *Saf. Sci.* 47, 459.
- Hokstad, P., Vatn, J., Aven, T., Sørum, M., 2004. Use of Risk Acceptance Criteria in Norwegian Offshore Industry: Dilemmas and Challenges. *Risk Decision and Policy* (July).
- Hopkins, A., 2008. Failure to Learn: the BP Texas City Refinery Disaster. CCH Australia, Reprint edition. July 1.
- Hopkins, A., 2009. Thinking about process safety indicators. *Saf. Sci.* 47 (4), 460–465.
- IOGP – The International Association of Oil and Gas Producers, 2016. Standardization of Barrier Definitions. Report 544, Supplement to Report 415.
- IOGP – The International Association of Oil and Gas Producers, 2018a. Asset Integrity – the Key to Managing Major Incident Risks. Report 415.
- IOGP – The International Association of Oil and Gas Producers, 2018b. Process Safety – Recommended Practise on Key Performance Indicators. Report 456.
- ISO 14224, 2016. Petroleum, Petrochemical and Natural Gas Industries — Collection and Exchange of Reliability and Maintenance Data for Equipment.
- ISO 17776, 2016. Petroleum and Natural Gas Industries — Offshore Production Installations — Major Accident Hazard Management during the Design of New Installations.
- ISO 41011, 2017. Facility Management — Vocabulary.
- Jansen, M., 2015. Shell maintenance excellence. In: *SPE-177964*. Presented at the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, 9-12 November 2015.
- Kaganski, S., Toomalu, S., 2017. Development of key performance selection index model. *Journal of Achievements in Materials and Manufacturing Engineering* 82 (1), 33–40.
- Kaplan, R.S., Norton, D.P., 1996. *The Balanced Scorecard: Translating Strategy into Action*, 1st. edition. Harvard Business Review Press, Boston, MA, USA.
- Kaufman, R., Oakley-Browne, H., Watkins, R., Leigh, D., 2003. *Strategic Planning for Success: Aligning People, Performance, and Payoffs*. Jossey-Bass/Pfeiffer (by John Wiley & Sons, Inc.), San Francisco, CA.
- Kavanagh, D., 2013. *Advantage – A Roadmap for Entrepreneurs and Leaders in the Digital Age*, Kindle edition. Published by Declan Kavanagh. <http://intelligentorg.com/wp-content/uploads/2013/11/@Note-10-SMART-squared.pdf>.
- Kongsvik, T., Johnsen, S.Å.K., Sklet, S., 2011. Safety climate and hydrocarbon leaks: an empirical contribution to the leading-lagging indicator discussion. *J. Loss Prev. Process. Ind.* 24 (4), 405–411.
- Lawlor, K.B., Hornyak, M.J., 2012. SMART goals: how the application of SMART goals can contribute to achievement of student leaning outcomes. *Developments in Business Simulation and Experiential Learning* 39, 259–267.
- Leveson, N., 2015. A systems approach to risk management through leading safety indicators. *Reliab. Eng. Syst. Saf.* 136, 17–34.
- Longhi, A.E.B., Pessoa, A.A., de Almada Garcia, P.A., 2015. Multiobjective optimization of strategies for operation and testing of low-demand safety instrumented systems using a generic algorithm and fault trees. *Reliab. Eng. Syst. Saf.* 142, 525–538.
- Matheson, D., Matheson, J., 1998. *The Smart Organization – Creating Value through Strategic R&D*. Harvard Business School Press, Boston, MA, USA.
- NORSOK D-010, 2013. *Well Integrity in Drilling and Well Operations*, fourth ed.
- OECD – Organisation for Economic Cooperation and Development, 2008. *Guidance on Developing Safety Performance Indicators Related to Chemical Accident Prevention, Preparedness and Response*, second ed. Series on chemical accidents no. 19.
- Øien, K., Utne, I.B., Herrera, I.A., 2011. Building safety indicators: Part 1 – theoretical foundation. *Saf. Sci.* 49 (2), 148–161, 2011.
- Parida, A., Chattopadhyay, G., 2007. Development of a multi-criteria hierarchal framework for maintenance performance measurement (MPM). *J. Qual. Mainten. Eng.* 13 (3), 241–258.
- Parida, A., Kumar, U., 2006. Maintenance performance measurement (MPM): issues and challenges. *J. Qual. Mainten. Eng.* 12 (3), 239–251.
- Payne, S.C., Bergman, M.E., Beus, J.M., Rodríguez, J.M., Henning, J.B., 2009. Safety climate: leading or lagging indicator of safety outcomes? *J. Loss Prev. Process. Ind.* 22 (6), 735–739.
- Peng, W., Sun, T., Rose, P., Li, T., 2007. A semi-automatic system with an iterative learning method for discovering the leading indicators in business processes. In: *Proceedings of the 2007 International Workshop on Domain Driven Data Mining, ACM, San Jose, CA, USA — August 12-15*.
- PSA – Petroleum Safety Authority Norway, 2013. *Principles for Barrier Management in the Process Industries*.
- RapidBI, 2016. *Writing SMARTer objectives and goals*. Uploaded. <https://rapidbi.com/writesmartobjectives/>. (Accessed 2 May 2016).
- Roos, C.J., Myers, P.E., 2015. *The Engineer's Guide to Overfill Prevention*. Emerson Process Management.
- Saaty, T.L., 1980. *The Analytic Hierarchy Process*. McGraw-Hill Inc.
- Saleh, J.H., Marais, K.B., Favaro, F.M., 2014a. System safety principles: a multidisciplinary engineering perspective. *J. Loss Prev. Process. Ind.* 29, 283–294.
- Saleh, J.H., Haga, R.A., Favaro, F.M., Bakolas, E., 2014b. Texas City refinery accident: case study in breakdown of defense-in-depth and violation of the safety-diagnosability principle in design. *Eng. Fail. Anal.* 36, 121–133.
- Selvik, J.T., Abrahamsen, E.B., 2015. A review of safety valve reliability using failure fraction information. In: *Proceedings of the 25th European Safety and Reliability Conference, Zurich, Switzerland, pp. 7–10, September 2015*.
- Selvik, J.T., Stanley, I., Abrahamsen, E.B., 2020. SMART criteria for quality assessment of key performance indicators used in the oil and gas Industry. *Int. J. Perform. Eng.* 16 (7), 999–1007.
- Smart, K., Blakey, K., 2014. Achieving maintenance excellence in Maersk oil Qatar. IPTC 17623. In: *Proceedings of the International Petroleum Technology Conference (IPTC) Doha, Qatar, pp. 20–22, January 2014*.
- Smith, R., Mobley, R.K., 2008. Chapter 6 – Key Performance Indicators. *Rules of Thumb for Maintenance and Reliability Engineers*. Butterworth-Heinemann, pp. 89–106.
- Summers, A.E., Hearn, W., 2010. Overfill protective systems—complex problem, simple solution. *Process Saf. Prog.* 29 (4), 283–287.
- Swuste, P., Theunissen, J., Schmitz, P., Reniers, G., Blokland, P., 2016. Process safety indicators, a review of literature. *J. Loss Prev. Process. Ind.* 40, 162–173.
- Tamim, N., Laboureur, D.M., Mentzer, R.A., Hasan, A.R., Mannan, M.S., 2017. A framework for developing leading indicators for offshore drillwell blowout incidents. *Process Saf. Environ. Protect.* 106, 256–262.
- The BP US Refineries Independent Safety Review Panel, 2007. *The report of the BP US refineries independent safety review panel*. Downloaded 7 August 2020, from: <http://sunnyday.mit.edu/Baker-panel-report.pdf>.
- Vinnem, J.E., 2012. On the analysis of hydrocarbon leaks in the Norwegian offshore industry. *J. Loss Prev. Process. Ind.* 25 (4), 709–717.
- Vukomanovic, M., Radujkovic, M., 2013. The balanced scorecard and EFQM working together in a performance management framework in construction industry. *J. Civ. Eng. Manag.* 19 (5), 683–695.
- Waite, P., 2013. Recurring accidents: overfilling vessels. *Chem. Eng. Lond.* 861, 40–44.
- Wood, D.A., 2016. Asset portfolio multi-objective optimization tools provide insight to value, risk and strategy for gas and oil decision makers. *J. Nat. Gas Sci. Eng.* 33, 196–216.