Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

Review

# Near-miss management systems and observability-in-depth: Handling safety incidents and accident precursors in light of safety principles

Maria Grazia Gnoni [a], Joseph Homer Saleh [b,]*

[a] *Department of Innovation Engineering, University of Salento, Lecce, Italy*
[b] *School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA*

ARTICLE INFO

*Article history:*
Received 4 August 2016
Accepted 15 August 2016

*Keywords:*
Near-miss
Safety management system
Observability-in-depth
Safety principles
Accident precursors

ABSTRACT

Accident precursors and near-miss management systems (NMS) are important safety tools in industries with major accident hazards, such as the airline industry, the nuclear industry, and the chemical industry, and they are increasingly adopted in other sectors such as manufacturing, construction, and health care. The aim of a NMS is to "harvest value" from near-miss data by assessing and prioritizing their risk implications, identifying their failure generating mechanisms, and guiding interventions and safety improvements and awareness. Recognizing that learning from near-misses is less costly than learning from accidents, the main value of a NMS is in the learning loop it provides within and across organizations, in focusing safety resources on addressing unsafe acts, reducing unsafe conditions and procedures, and improving design and operational safety issues. The present study first provides an updated review and synthesis of key ideas and challenges of NMS. It then proposes and examines important synergies between fundamental safety principles adopted in risk management, including defense- and observability- in depth, and NMS. Safety principles offer a new lens by which to view NMS. One important result is that near-miss data can be classified and interpreted in light of safety principles violated, and that safety interventions can be particularly effective when organized around such findings, the objectives being to (re-)establish and strengthen compliance with safety principles through workforce training, system redesign, and/or improved operational procedures. Finally, it is argued that NMS is one of the pillars of the implementation of observability-in-depth, and that the boundaries with the two other pillars (fault detection/online monitoring, and inspection) are likely to be blurred in the future, and that the next generation NMS (2.0) will likely integrate data from multiple sources to improve the efficacy of precursor identification, prioritization, and safety interventions, and ultimately accident prevention.

**Contents**

* Corresponding author.
*E-mail address:* jsaleh@gatech.edu (J.H. Saleh).

http://dx.doi.org/10.1016/j.ssci.2016.08.012
0925-7535/

## 1. Introduction

This work brings together two strands in the literature on accident prevention and system safety: the development and formalization of general, domain-independent, system safety principles on the one hand, and the examination of near-miss management systems on the other hand. The objectives of the present work are threefold: (i) to review and synthesize the key ideas and challenges of near-miss management systems; (ii) to introduce a set of high-level, domain-independent, safety principles to a broader audience, especially the readership and safety professionals involved in near-miss management systems; and (iii) to examine the relation and synergies between safety principles and near-miss management systems, in particular the observability-in-depth principle, and to explore how safety principles can help inform the design and operation of near-miss management systems and improve their effectiveness.

### *1.1. Definitions*

Before delving into the topic, it is worth clarifying the two related terms of "accident precursor" and "near-miss". Although no generally agreed upon definitions are available, different authors adopt slightly different interpretations of these concepts, their general meaning is intuitive and easily understandable. For example, the National Academy of Engineering defines accident precursors as "conditions, events, and sequences that precede and [can] lead up to an accident" (NAE, 2004), and NASA defines an accident precursor as "an anomaly [off-nominal occurrence or condition] that signals the potential for more severe consequences that may occur in the future, due to causes that are discernible from its occurrence today" (NASA, 2011). And the U.S. Nuclear Regulatory Commission defines an accident precursor as "an observed event and/or condition at a plant, [which] when combined with one or more postulated events (e.g., equipment failures, human errors) could result in core damage" (NRC, 2008). An accident precursor is best understood in relation to the notion of accident sequence, or the sequence of events starting from an off-nominal initiating event, followed by increasingly more hazardous events/states, and leading up to an accident—the uncontrolled release of energy and its adverse consequences (e.g., injuries and loss of life, destruction of property or infrastructure, environmental damage).

An accident precursor can thus be conceived of as any truncation of an accident sequence. Moreover, the precursor can be qualified by its closeness to the complete accident sequence. According to this point of view, a near-miss is a special type of accident precursor for which the truncation of a complete accident sequence is minimal (close to the accident end-state or occurrence). In other words, a near-miss is very similar to an accident sequence with the exception of a few missing elements or ingredients, which translate into a few missing events (truncation) in the accident sequence (Saleh et al., 2013). The further an accident sequence advances before it is interrupted, the more hazardous the situation is, and the more appropriate is its characterization as a near-miss (sometimes referred to casually as a "close call").

In short, a near-miss has many of the ingredients (conditions) and generating mechanisms (causal factors) of an accident sequence with the exception of a few missing ones, which prevent it for further escalating and leading to the accident and its dire consequences. Since near-misses and accidents share some/many common causes, learning from the former (near-misses) and eliminating their causes makes a positive contribution toward preventing the latter (accidents).

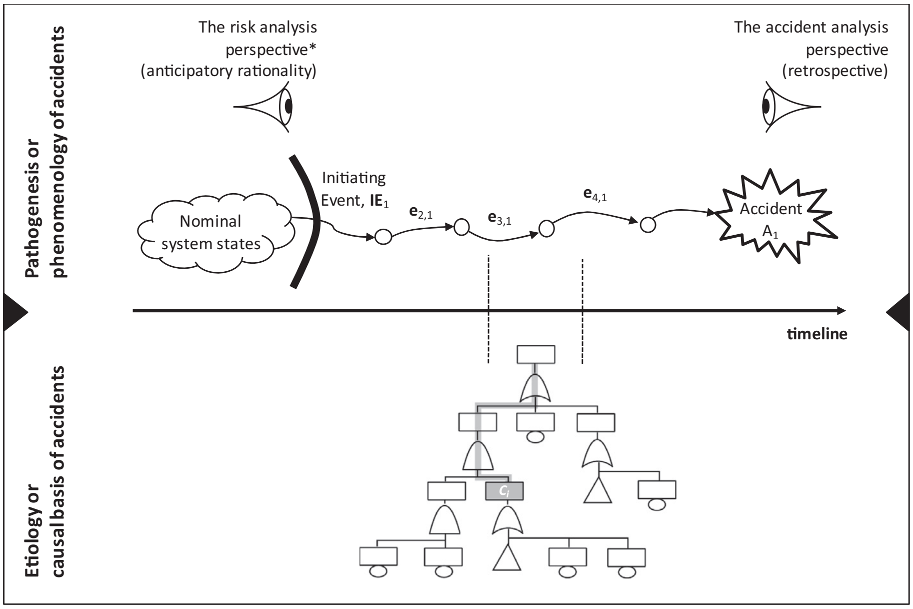
As a side note, the severity of consequences is often used as a distinguishing feature between different terms, a disaster for example on one end of the spectrum involves many casualties (severe consequences), and an incident on the other end of the spectrum involving few or none (light adverse consequences). In this view, an accident precursor or near-miss is further out beyond a safety incident on this severity spectrum. This is the aspect favored by the NASA definition, "an anomaly [off-nominal occurrence or condition] that signals the potential for more severe consequences that may occur in the future."

It is worth pointing out that some level of ambiguity exists in the definitions of accident precursors and near-misses, and that the distinction between the two is to some extent subjective (e.g., where to draw the line between one and the other, and with other terms such as "incident"). This semantic wiggle room however is not detrimental to a proper understanding of these terms. When sharper definitions are need for some specific purpose and within a particular organization or context, more ad hoc nuances can be added.

One additional term that is usually subsumed under the definition of accident precursor but deserves some special attention is that of an accident pathogen. An accident pathogen is an adverse latent or pre-existing condition, passive or with no impact on the system output until activated or triggered by other adverse occurrences (see Fig. 1). When compounded with other factors, an (activated) accident pathogen can further advance an accident sequence, precipitate an accident, or aggravate its consequences (Bakolas and Saleh, 2011). For example, a failed emergency power system is an accident pathogen at a nuclear power plant: should the main power system fail, this latent adverse condition will precipitate the accident, or it will cause the sequence to further advance toward a core meltdown (Saleh et al., 2013). In Fig. 1, $C_i$ is an example of an accident pathogen when this accident sequence has not been triggered yet or when it has stopped at the event $e_{3,1}$; the AND gate above $C_i$ does not have the two causes active to allow that the failure logic to propagate up the tree.

### *1.2. Learning loops and near-miss management systems*

Near-miss management systems (NMS), also referred to by other terms such as accident precursor programs, or incident (safety) reporting systems, are concerned with the broadest definition of near-misses and include adverse conditions (accident pathogens), unsafe acts and procedures, and adverse events or sequences of events "that precede and [can] lead up to an accident". All these aspects constitute an important source of knowledge when their safety implications are properly understood. The aim is to learn from collected data about near-misses (broadly


The risk analysis perspective* (anticipatory rationality)
The accident analysis perspective (retrospective)
Pathogenesis or phenomenology of accidents
Nominal system states
Initiating Event, IE1
e2,1
e3,1
e4,1
Accident A1
timeline
Etiology or causal basis of accidents
Cj

*Multiple accident trajectories are posited; only one is shown here for readability

**Fig. 1.** Accident sequence (phenomenology) and causal basis (etiology) of accidents.

defined), to understand their risk implications, and to transform that data into knowledge for guiding safety interventions and contributing to accident prevention. **Learning from near-misses is less costly than learning from their fully developed more destructive analogs, namely accidents.**

Since "learning" is intimately related to near-miss management systems (their *raison d'être*), it is worth briefly noting what this concept entails. At the individual level, learning can be loosely defined as the acquisition of knowledge or skill and the modification of behavior due to (the understanding of) previous experience (Merriam-Webster). The modification part is important for our purposes. According to Sterman (1994), "learning is a feedback process in which our decisions alter the real world, we receive information feedback about the world and revise the decisions we make and the mental models that motivate those decisions." Beyond the individual level, learning can also occur at the organization level, the broader industry level, and the regulatory level (Hovden et al., 2011). Near-miss management systems are an essential part of the operational feedback process both within and across organizations, their purpose is to collect and interpret near-miss operational data, to transform said data into information and actionable knowledge, and to trigger safety intervention for the modification of unsafe conditions in the design and operations of systems, including unsafe acts and procedures, to prevent their re-emergence with potentially more serious consequences. Phimister et al. (2003) rightfully identified that **the purpose of near-miss management systems is "to harvest value from near-miss [data]"**; the process by which this harvesting can take place, and the different value streams that ensue are examined hereafter. The remainder of this work is organized as follows: Section 2 reviews and synthesizes key ideas and challenges in the design and operation of near-miss management systems based on field experience from various hazardous industries. Section 3 discusses system safety principles. Section 4 brings these two elements together and explores the relation and synergies between system safety principles and near-miss management systems, in particular the observability-in-depth principle, and it examines how safety principles can help inform the design and operation of near-miss management systems and improve their effectiveness. Section 5 concludes this work.

## 2. Near-miss management systems: review, key ideas, and challenges

Following both accidents of the space shuttles Challenger and Columbia, the investigations noted that repeated precursors were present in both cases, and that "processes in place at the time did not respond [...] in a way that shed light on their true risk implications (NASA, 2011). Similar situations were found following the Three Mile Island nuclear accident (Hopkins, 2001), the Texas City refinery accident (Saleh et al., 2014a), and score of others cases. These instances fundamentally reflect a failure of the Near-miss Management system (NMS) or accident precursor program[1] (used interchangeably hereafter) within and across organizations, or the unavailability of such systems.[2] What are NMS? How are they structured? And what are the key issues and challenges associated with their design and operations? We briefly discuss these questions in this section.

NMS consist in an organizational structure and function with people, processes, and IT support or infrastructure, and whose objective is to collect and prioritize anomaly and precursor data, to interpret and assess their risk implications, and to transform this data into risk-informed interventions and safety improvements and awareness. Its end objective is to help improve accident prevention and sustain safety in a variety of ways (technical, operational, and organizational). The system teases out the failure generating mechanisms in, and the risk implications of, anomaly and precursor data and reflects them back to the organization in a variety of ways to have them addressed.

The design and operation of NMS varies across industries (Phimister et al., 2003) (e.g., for a manufacturing company or an airline operator), and depending on whether it is implemented within a company or at the regulatory level overseeing an entire industry (e.g., the U.S. Nuclear Regulatory Commission or the

[1] Also known as incident reporting system in the healthcare industry.

[2] Following the Columbia accident, the accident investigation board found "no process for regularly tracking and assessing the performance of [...] the agency [NASA] leading indicators and safety information system", and recommended the establishment of "an aggressive accident sequence precursor (ASP) [program/team] to identify accident precursors for all remaining shuttle flights...".

Federal Aviation Administration), some differences in objectives and implementations are expected. These issues are discussed next.

### 2.1. Objectives of NMS

The design and implementation of a NMS is intimately related to its objectives. These are often taken for granted and are unstated, but it is worth explicitly articulating them. Several objectives for NMS are found in the literature. For example, Van der Schaaf et al. (1991) argues for three general objectives of NMS:

(i) To gain a **qualitative insight** into how accident precursors can develop into accidents.
(ii) To arrive at a "statistically reliable **quantitative insight** into the [frequency of] occurrence of factors or combination of factors [adverse conditions, precursors] giving rise to accidents".
(iii) To **sustain safety vigilance** within an organization, that is, "to maintain a certain level of alertness to danger, especially when rates of injuries and accidents are low".

Some (minor) variations on these objectives are found in the literature. For example, the NASA accident precursor analysis program (NASA, 2011) sets the following objectives:

(i) To evaluate observed anomaly and precursor data and determine if the mechanism at their origin could recur with more severe consequences under different circumstances.
(ii) To find in precursor data underappreciated risks, which could result in severe consequences under different circumstances.
(iii) To leverage anomaly and precursor data to better predict risk and gauge the likelihood of losses (this is similar to the quantitative insight objective of Van der Schaaf); also stated as "to determine the safety implications of [precursors] and reflecting them back into the system's risk model".

These objectives reflect different aspects of the broader "learning from precursors[3]" heading; they are more specific about the different types of learning than should occur. At the industry or regulatory level, when precursor data is collected (companies may be mandated by law to provide such data to the regulators), additional objectives can be pursued when the data is aggregated from an entire sector not just a single plant, such as analyzing safety trends across the industry and "transferring the learning to other organizations [from where the near-miss data was collected]" (Cambraia et al., 2010). For example, the U.S Nuclear Regulatory Commission has operated an Accident Sequence Precursor (ASP) Program since 1979. The primary objective of ASP is "to evaluate the risk from all operating nuclear power plants (not individual plants), [...], and this has significantly influenced the way the analyses are done, the nature of the results, and the types of insights expected from the program" (Sattison, 2003). The ASP Program is used to monitor performance against safety goals across the entire industry established by the Commission (NRC, 2008). More specifically, the official objectives of the ASP Program are as follows: (1) "To provide a comprehensive, risk-informed view of nuclear power plant operational experience and a measure for trending nuclear power plant core damage risk [across the industry]; (2) to provide a partial check on dominant core damage scenarios predicted by probabilistic risk assessments [that is, input to improve risk models]; and (3) to provide feedback to regulatory activities." Similarly, the European Union (EU) has mandated that Member States report near-misses that occurred at establishments falling under the Seveso Directive, that is, the legislation covering major accident hazards in establishments containing large quantities of dangerous substances such as chemicals (Nivolianitou et al., 2006). The centralized database of these near-misses is known as the Major Accident Reporting System (MARS) and it is meant to facilitate the exchange of lessons learned involving dangerous substances across member states, to analyze safety trends within the industries covered across the European Union, and ultimately to help improve accident prevention.

The objectives for NMS discussed so far reflect an *analysis* perspective. Their importance cannot be understated, but they are also limited and lack an *action-oriented component* and *intervention* mindset. A NMS cannot be confined to analyzing data; it has to lead to safety interventions that address the mechanisms that generated the accident precursors (whether the adverse conditions, behaviors, or events), disseminating the information, and monitoring the effectiveness of the safety solution(s) implemented. In short, the end-objective of a near-miss management system should prominently include, "to improve safety performance" within or across organizations. These objectives will be reflected in the framework and processes of a near-miss management system discussed next.

### 2.2. Framework for designing a NMS, and key issues and challenges

NMS are broadly applied in hazardous industrial sectors from the nuclear industry to chemical and oil & gas sectors. The airline industry also has a long tradition of using NMS. Several applications are recently developing in new sectors such as construction (Goldenhar et al., 2003; Wu et al., 2010) and manufacturing (Gnoni et al., 2013). However, no established standard for the design and implementation of NMS exists, and it is common to find such systems with different features and phases across industries. One typical framework for an effective NMS was proposed by Van der Schaaf (1992, 2013) and it includes the following steps:

- *Step 1:* Near-miss reporting and identification.
- *Step 2:* Prioritization and selection of near-miss events for further analysis.
- *Step 3:* Root cause analysis and risk reduction measures.
- *Step 4:* Follow up measures.

Phimister et al. (2003) proposed a similar framework shown in Fig. 3.

The identification step reflects the realization that a safety incident has occurred (a near-miss, broadly defined) or an adverse condition in the system is identified. A safety issue, the authors argue, identified but not reported is of little use or learning value, except perhaps to the individual who stumbled upon it or experienced it. The next step in Fig. 3 involves screening and prioritizing of the data collected (according to some metrics, which will be discussed shortly). The prioritization phase is important, especially when extensive anomaly and near-miss data are collected, to properly allocate limited resources for the subsequent safety analysis and interventions where they are most needed. The next step involves careful safety analysis of the remaining data, to identify the underlying causal mechanisms generating these anomalies and adverse conditions, a sort of why-because analysis, applied repeatedly or recursively to the data set. In this step, the main locus of "learning", **the near-miss data is transformed into safety information and knowledge**. In this phase, **a prospective analysis is also conducted of hypothetical futures extrapolating the precursors to accidents that could have happened, to better appreciate their safety implications and prioritize the need for intervention and safety solutions**. This is followed by a phase during which corrective actions are devised, that is, solutions to

[3] Broadly defined to include anomalies, adverse conditions, accident pathogens, and near-misses.

address, eliminate, or mitigate these anomalies and their generating mechanisms. The last two phases in this framework involve disseminating the corrective actions to relevant parties, implementing these solutions, and monitoring their effectiveness (along with other follow-up, as needed). Each step in this framework raises a set of issues and challenges. These are discussed next.

#### *2.2.1. Key issues and challenges with the identification and reporting phases of a NMS*

Reporting of near-miss data can be done in a variety of ways. For example, they can be reported by "the person experiencing the near-miss on either a voluntary or mandatory basis" (Van der Schaaf, 1992), or they can be reported by observational methods designed specifically to scan for off-nominal adverse events or conditions, either by designated individuals or automated sensors. This is the case for example of routine checks of aircraft black boxes to detect excursions of parameters outside the flight envelope (this approach has a significant untapped potential for other industries, and it is worth carefully exploring in the nuclear, the chemical, and the oil and gas industries, to mention a few). When individuals are reporting near-misses, the process can be designed to be conducted anonymously or not; the former is meant to encourage reporting whereas the latter raises the specter of liabilities, which constitute a strong disincentive to reporting. When the individual is named or identified in the near-miss report, follow-up is possible and a richer set of contextual information may be collected to better appreciate the safety implications of the near-miss. The process in this case can include granting immunity to the individual (under certain conditions, again to encourage reporting). It is important to avoid a punitive mindset when designing and operating a NMS, "to foster a rich reporting culture" and trust in the system (Barach and Small, 2000), and to promote a proper learning and safety culture in the organization. Other disincentives to reporting near-misses can be found in Van der Schaaf and Kanse (2004). Two additional important challenges have been noted in the literature regarding the reporting of near-miss data: (1) confusion regarding what is reportable; and (2) the extent of data collected and whether report rates are a desirable metric to decrease over time.

(1) **Confusion regarding what is reportable to the NMS:** several reports indicate that employees are often confused as to what constitutes a near-miss (or precursor), and consequently what should be reported. In an extensive study of near-miss systems in the chemical industry, Phimister et al. (2003) found that 68% of respondents expressed such confusion. It is thus advisable to take this issue seriously and address it upfront when a NMS is being designed or fielded. The essence of the confusion, we believe, lies in the understanding of a near-miss as solely an event-based concept, but not an existing adverse condition (accident pathogen). For example, "a crane operator accidentally drops a [heavy] load" (Van der Schaaf et al., 1991) before it reaches its destination on a construction site. Chance had it that no one was underneath the load path at the time, and as a result no fatalities occurred. This is an event-based incident, and as such no confusion would result whether this qualifies as a near-miss and that it should be reported.[4] However, when confronted with an adverse condition (not event), for example, a loose scaffolding on a construction site, a heavily corroded safety valve in a chemical plant, or rust on the external surface of an aircraft (sometimes an indication of dangerous cracks in the structure), employers are less likely to report such instances to a NMS. This situation is further aggravated by event-based definitions of near-misses found in the literature. For example, Cambraia et al. (2010) specified that for their system (and consequently what is reportable), a near-miss is "an instantaneous event, which involved the release of energy and had the potential to generate an accident." We believe this is a flawed approach to the use of NMS (for reasons discussed in Section 1). While we recognize that different industries can, and should, have different definitions of near-misses tailored to their particular circumstances, **a NMS should not be confined to event-based safety incidents; many safety blind spots would remain and several learning opportunities to improve safety would be forfeited by doing so**. As noted in the Introduction, we strongly believe NMS should be concerned with the broadest definition of near-misses and include adverse conditions (accident pathogens), unsafe acts and procedures, and adverse events or sequences of events "that precede and [can] lead up to an accident". It is also important that this definition and scope of NMS be clearly communicated to the employees. There are of course downsides to this broad definition of near-miss and the operation of NMS, and these are acknowledged next.

(2) **Extent of near-miss data collected:** By examining the second issue, the literature on the subject indicates that the average number of near-miss reported in the oil & gas, the chemical, and the construction industries, varies from 0.5 to 4 per FTE per year (Jones et al., 1999; Phimister et al., 2003; Cambraia et al., 2010). With a medium to large size organization, several hundreds or thousand reports of near-misses might be filed every year. This might be considered an instance of data overload if the organization does not have the right processes and filters in place to deal with this volume of data. But it can also be considered a rich data set to mine and use for guiding safety improvements and awareness (and reflecting in part the safety conditions and safety culture in the organization). Restricting this data set at the collection stage might lead to unnecessary safety blind spots. The proper filters for interpreting and assessing the safety implications of the data set are better placed downstream in the process, as shown in Fig. 3. **An important quality of a good NMS is that it ought to minimize both "missed detections" of accident precursors as well as "false positives"**; the latter would clog the system and divert resources from where they are needed most. This situation is illustrated in Fig. 3. How to achieve this dual objective is contingent on the specifics of the industry or plant under consideration and requires skills and a careful balancing act on behalf of the safety professionals designing and operating said NMS. This includes among other things a clear and unambiguous definition of what ought to be reported, a proper dissemination of the scope of the near-miss program, and a good risk awareness and safety training of the employees. Another closely related issue with the extent of near-miss data reported is whether the program should set as an objective the reduction or increase of the number of near-miss reported. There are too many confounding factors to make a general answer (statistically) meaningful, and as such, the number of reported near-misses cannot be credibly used as a positive or negative indication of safety.[5] Manage-

[4] There are several possible learning opportunities in this near-miss case, for example, reviewing the design of the crane grips, improving operation and operator training (e.g., limiting the acceptable weights that can be handled), and/or establishing a safety policy/procedure that workers never stand underneath the path of load being carried by a crane.

[5] In Fig. 3, the number of reported near-misses is noted as $\hat{n}_p$ and the actual number of (unknown) near-misses is $n_p$. Seeking to minimize $\hat{n}_p$ when the relationship between $\|n_p - \hat{n}_p\|$ is not understood is a flawed objective.

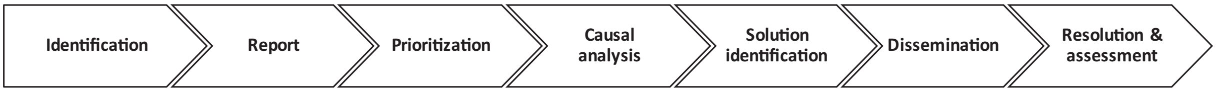


**Fig. 2.** Near-miss management framework and steps (proposed by Phimister et al. (2003)).

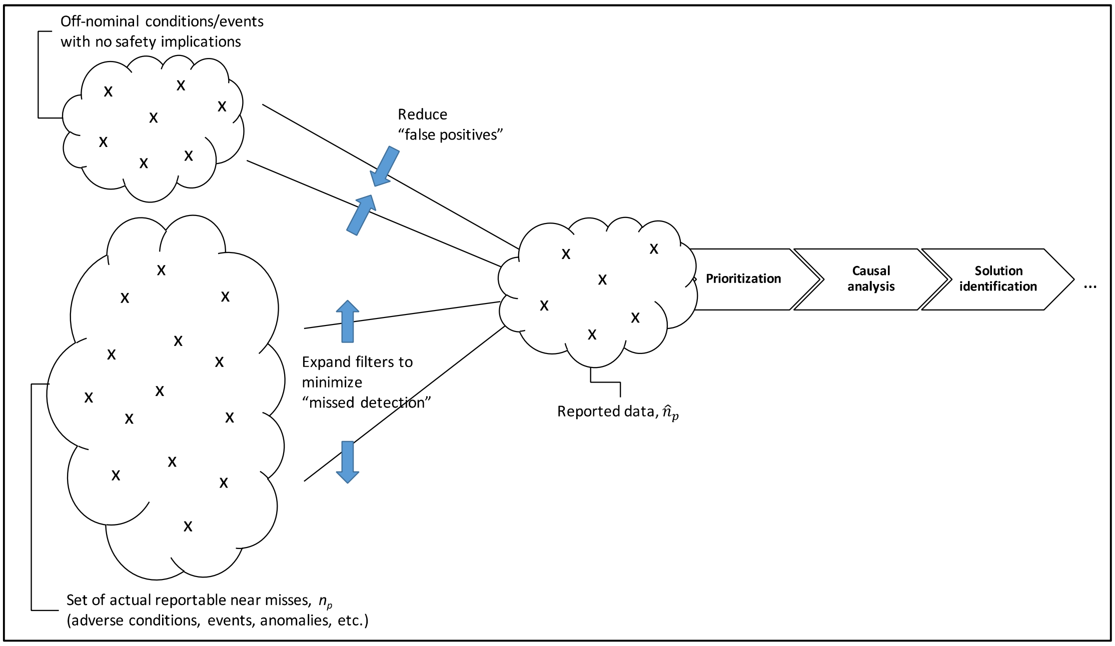

Off-nominal conditions/events with no safety implications
Reduce "false positives"
Expand filters to minimize "missed detection"
Prioritization
Causal analysis
Solution identification
...
Reported data, $\hat{n}_p$
Set of actual reportable near misses, $n_p$ (adverse conditions, events, anomalies, etc.)

**Fig. 3.** Illustration of the dual objective of a NMS: to minimize both miss detections of accident precursors and false positives.

ment should refrain from indicating that a safety objective is to compress the number of reported near-misses—that is the wrong target (emphasis on *reported*). The objective should be to always report what ought to be reported (and provide clear training to support that). Too high or too low of a number of near-misses reported might be indicative of pathological situations within the company or the plant, but the effectiveness of a near-miss management system (and safety interventions), if needed, ought to be carefully planned and conducted, like with any intervention analysis, and it cannot rely on vague intuition and basic descriptive statistics such as the number of near-miss reports filed.[6]

### *2.2.2. Key issues and challenges with the down-selection and prioritization phase of a NMS*

Several studies have confirmed trends exhibited in a typical safety pyramid shown in Fig. 4, namely that for every accident, two or three orders of magnitude more near-misses exist or might be lurking (Bird and Germain, 1996; Masimore, 2007; Manuele, 2011). Furthermore, assuming a rate of 1 near-miss report per FTE per year, a medium to large size organization will collect several hundreds or thousands of near-miss reports every year. It is clear therefore that some down-selection and prioritization is needed to carefully allocate an organization's limited resources for safety considerations. Not all reported near-misses can proceed through the system in Fig. 2 and undergo further detailed risk analysis (and possibly lead to safety interventions).

The two related questions are: which ones to focus on? And how to go about this prioritization to minimize the likelihood of dismissing important near-misses and maximize the likelihood of filtering out the irrelevant ones? This is obviously a critical phase in any NMS, and it ought to be tailored to the specific industry under consideration. There is no one-size-fits-all and no standard approach to this issue. For example, in the U.S. nuclear industry, the NRC mandates the reporting of near-misses by nuclear power plants. The reports are known as Licensee Event Reports (LER), and they undergo a two-step review process to prioritize them, which in this case is a binary decision of rejection or acceptance for further analysis. The LER are assessed against a set of specific screening criteria by two engineers (the details can be found in NRC (2008)). Furthermore, the NRC classifies as a high-priority *significant* precursor an LER that result in a "condition core damage

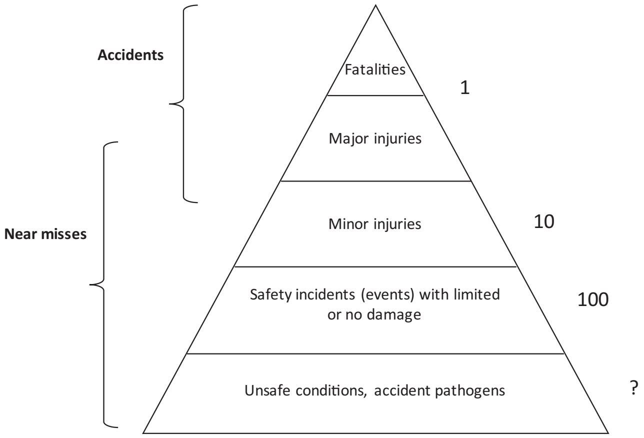

Accidents
Near misses
Fatalities
1
Major injuries
Minor injuries
10
Safety incidents (events) with limited or no damage
100
Unsafe conditions, accident pathogens
?

**Fig. 4.** A typical safety pyramid – initially proposed by Heinrich (1931) and updated by Bird and Germain (1996), Masimore (2007), Manuele (2011) and others.

[6] Incidentally the tools and techniques of Machine Learning offer great potential for exploring near-miss data and becoming an integral part of NMS. This topic is left as a fruitful venue for future work.

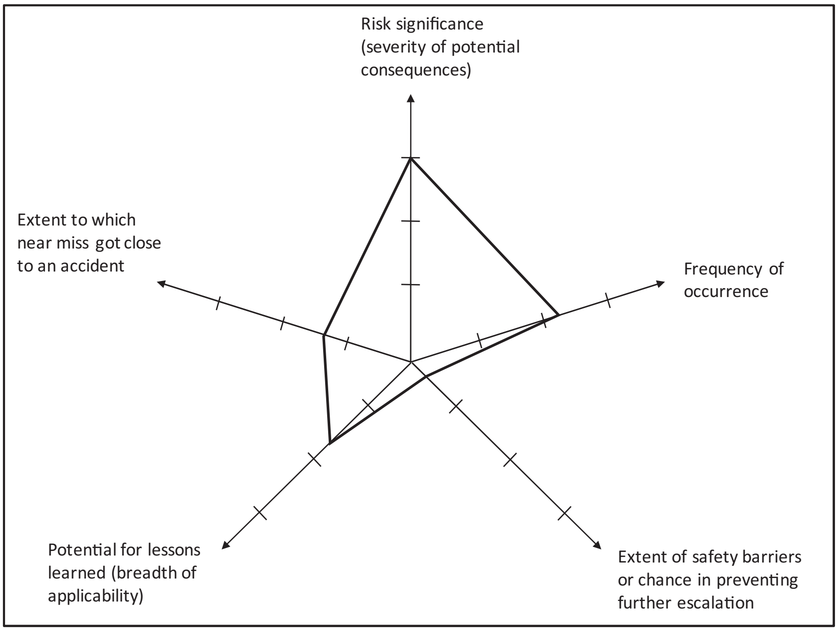


**Fig. 5.** Illustration of a radar plot for a quick qualitative assessment of near-miss reports, and down-selection for further more detailed analysis or rejection (scales can be rough Low–Medium–High categories).

probability (CCDP) or an [increase] in core damage probability ΔCCDP of $10^{-3}$". The relevant point for our discussion is that near-miss data are prioritized based on their risk significance, and that the results of the prioritization, especially when with an acceptance/rejection filter, are better cross-checked for confirmation. More generally, the prioritization would benefit from including several filters and (qualitative) considerations:

(1) An assessment of the risk significance or how severe the consequences would have been had the precursor escalated to a full fledged accident.
(2) The extent to which the near-miss or accident sequence has advanced before it was blocked or terminated, that is, how close it came to the accident. This would measure a sort of "distance" (or norm in a mathematical sense) between the near-miss and the accident end-state.
(3) The frequency of occurrence of a near-miss, and whether it amounts to a trend of similar adverse occurrences or conditions (repeated offenses).

Additional prioritization filters can include for example:

(4) Near-misses for which there is little or no protection, and precursors that did not further escalate to a full fledged accident by chance, not because of proper safety barriers.
(5) Near-miss for which "the potential of *lessons learned* is farther reaching" than the scope of the particular report, and addressing it would be broadly beneficial (Phimister et al., 2003).

The prioritization phase is a critical link in a NMS. It is meant to conduct the equivalent of a medical of triage[7] and select from the near-miss reports, especially when a large number is filed, which ones ought to proceed through the NMS (Fig. 2) and be subjected to more careful and detailed risk analysis. This is a quick, qualitative, down-selection process. This initial filtering is best conducted by individuals who are well versed in risk and safety issues within the industry of interest. Several authors have examined the use risk matrices to prioritize near-misses (Ritwik, 2002; Cambraia et al., 2010; Gnoni and Lettera, 2012; Kleindorfer et al., 2012). This is simply a visualization tool when two (or more) criteria are adopted (e.g., traditionally likelihood and severity). It is important to recognize that there are several criteria by which near-misses can be prioritized (see previous 1–5 list). Sattison (2003) recognized that rankings can differ when a single criterion is selected. It is advisable not to reduce the complexity of the down-selection problem by adopting a single criterion, and to grapple instead with an overall holistic view of the different characteristics of the near-misses reported before selecting them for further analysis or rejecting them upfront.[8] A radar plot can be a useful visualization tool in this respect (see Fig. 5).

Like medical triage, the prioritization phase in a NMS is done rather quickly and qualitatively, and often when little information and much uncertainty about the reported near-misses is involved. In other words, the basis for assessing the reports against the considerations proposed earlier is likely to be lacking at this point, and will be further strengthened in the next *causal analysis* phase.

### *2.2.3. Key issues and challenges with the causal analysis phase of a NMS*

The *causal analysis* phase builds on the earlier *prioritization* phase, and it can start by revisiting the previous analyses albeit with significantly more depth and thoroughness for the (much) smaller number of precursors selected to proceed further through the NMS (Fig. 2). For example, assume the repeated foam strikes on the space Shuttle have been reported to a NASA accident precursor program (before the Columbia accident), and it made it past the *prioritization* filters. How can this report be handled? It is important first to recognize that two different types of analyzes and mindsets are necessary in this phase, otherwise the NMS is myopic if it doesn't include these two approaches:

[7] "The process of deciding which patients should be treated first based on how sick or seriously injured they are" (Merriam-Webster).

[8] The famous quip comes to mind, that "complex problems have simple, easy to understand, wrong answers". The designers of NMS should avoid this pitfall.

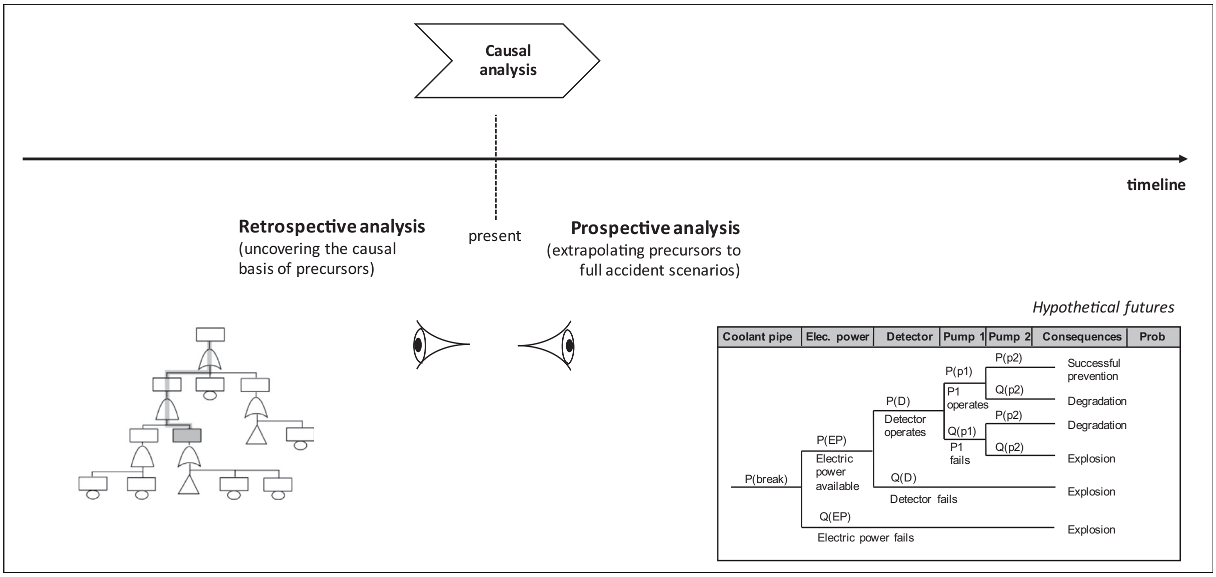


**Fig. 6.** Retrospective and prospective analyses of accident precursors.

(1) A retrospective analysis focused on identifying the causal basis of the precursor and its generating mechanism (more on this shortly).

(2) A prospective analysis of hypothetical futures extrapolating the precursors to accidents that could have happened.[9] That is, identifying different accident sequences and scenarios that could unfold by "passing through" the reported precursor(s) and to which a few additional ingredients are added. These additional ingredients have to be identified, and the reason why they were not present or did not occur (e.g., by chance or by design of effective safety barriers) has to be assessed. This is in essence risk analysis or the "imagination of failure" when a few elements of the accident scenario have already occurred. This is akin to reconstructing a(n accident) DNA when only few strands are available.

Fig. 6 shows that for the retrospective analysis, Fault Tree Analysis (FTA) can be a useful tool for structuring the examination of the causal basis of precursors (other tools can also be helpful). In simple terms, this examination consists in a repeated application of a Why-Because-Analysis (WBA) until factors of different nature (technical and organizational, including design, operation, and procedures) and operating over different time scales are identified. The latter is referred to as the temporal depth of causality of system accidents (Cowlagi and Saleh, 2013), and it is useful to acknowledge it in the case of near-misses as well when applicable. It is important to recognize that the prominence of a factor in the causal chain leading to an accident or near-miss is NOT determined by its temporal vicinity to the adverse event. That is, a *direct cause*[10] of a near-miss for example is not necessarily more important that other factors further upstream the causal chain (e.g., procedural or organizational). This will have important consequences for the *solution identification* phase.

The prioritization filters or considerations should be examined more carefully in this phase, and the safety analysts managing this process would benefit from having a checklist or a template, similar for example to what accident investigators have, to guide the examination of the causal basis of the near-miss. These can include for example factors related to:

- The design of the system or particular subsystem related to the near-miss (hardware and software).
- The training of the operators or technicians involved in the near-miss (liveware).
- The maintenance procedures and executions.
- The environmental factors involved (broadly defined).
- The managerial/organizational factors.

The list is not meant to be exhaustive, and specific checklist or template can be developed and tailored to a particular industry, for systematically and consistently going through the *causal analysis* phase. Finally, we acknowledge that in some cases, the previous suggestions might be an overkill, and a simpler, more straightforward examination of the causal basis of the near-miss may be sufficient.

In addition to the retrospective analysis, the prospective analysis is particularly important for properly understanding the risk implications of the near-miss, especially the (change in) likelihood and consequences of the postulated accident scenarios. This exercise consists in extrapolating the precursors to accidents that could have happened, and as such, Event Tree Analysis (ETA) can be a useful tool for structuring this examination of hypothetical futures (other tools can also be helpful). A new more informed prioritization of near-miss data can also take place at this stage following this prospective analysis, and as a result, it is useful to include a feedback loop to the previous *prioritization* phase to confirm, modify, or improve the use of the previous filters.

In both cases, it is important to clearly identify and understand (1) which safety barriers failed or were absent and enabled the occurrence of the near-miss, (2) which adverse events have little or no protection against their occurrence, and (3) whether the near-misses did not further escalate to a full fledged accident by chance, or because of proper safety barriers. This knowledge will feed into the *solution identification* phase discussed next.

### 2.2.4. Key issues and challenges with the solution identification phase of a NMS

Having understood the causal basis of the precursors (ingredients, generating mechanisms, and contributing factors)

[9] It can be argued that the prospective analysis can be done selectively on a few near-misses after the retrospective analysis is conducted and some serious flags are raised. We err on the side of caution with our suggestion that these two analyses be conducted on all selected near-miss reports that made it past the prioritization phase, if resources are available.

[10] Immediate cause preceding an event; leading to the event without any further intervention or contributing factors.

and how they could have further escalated into accidents, the next step switches to a designer mindset instead of the previous analyst mindset. It requires creativity in identifying ways to: (1) address and eliminate the generating mechanisms of the near-misses; and (2) ensure that safety barriers are in place and reinforced to consistently prevent further escalation. The military concept of a **kill chain** is useful here to appreciate the key issues in this *solution identification* phase: the objective of this phase is to identify the most efficient way to kill the near-miss generating mechanism, to eliminate (to the extent possible) its key ingredients for all future times. The theme of preventing recurrence of accident and near-misses is pervasive in the safety literature, and it is fundamentally anchored in this phase of a NMS. It goes without saying that this phase requires technical and operational ingenuity to identify where to intervene most (cost)effectively to prevent recurrence of the near-miss. Safety professionals are encouraged to avoid the temptation of quick *band-aid* solutions that do not address the fundamental generating mechanism of the near-miss. As noted in the NASA accident precursor analysis handbook, "the [near-miss or] off-nominal event is not simply resolved so that operations can continue; it is analyzed…and used to help understand and control risk in the future" (NASA, 2011).

### *2.2.5. Key issues and challenges with the dissemination and resolution & assessment phases of a NMS*

These last two phases in Fig. 3 are sometimes wrapped up and included in the previous *solution identification* phase. It is useful though to identify and acknowledge their functions: dissemination of the lessons learned from the near-miss to relevant parties, and the implementation of the corrective actions chosen. The last phase should also include provisions for monitoring the effectiveness of the implemented solutions or safety intervention; and to feed back that assessment to the *solution identification* team. The challenge in the *dissemination* phase is to identify the right parties to reach out to (in some cases, this can require training with new operational procedures and the target audience is clear) and to avoid information overload and over-dissemination (high frequency) of near-miss information and solutions.

Finally, the NMS should include a self-reflective provision for a periodic general assessment of its own performance (noted as *assessment* in Fig. 3). This can include for example an examination of the trends in near-miss data reported (by different covariates), an assessment of the quality of the *prioritization* phase in light of the subsequent results from the *causal analysis* phase, and a benchmarking of the effectiveness of the different solutions implemented (e.g., whether similar causal factors recur in later near-misses, and which approaches are more effective at eradicating specific near-miss generating mechanisms). This periodic assessment exercise can result in several feedback loops to all the previous phases in the NMS, to modify or improve upon them if needs be. This in a sense is a nested learning loop within a system designed fundamentally to learn from near-misses. Near-miss management systems should "walk their own talk" and have by design a self-assessment and learning feature to periodically evaluate and improve their own design and operation.

## 3. System safety principles

This section is a synthesis and abridgment of previous articles by the authors (Saleh et al., 2014a; Favaro and Saleh, 2014). It is included here to keep the present work self-contained and make it easier for the reader to follow through the subsequent developments without having to consult the other references. Additional details can be found in the original references.

In the previous section, we examined the key ideas and challenges of near-miss management systems. Recall that one of the objectives of the present work is to examine the synergies between system safety principles and near-miss management systems, and to explore how the former can help inform the design and operation of NMS and improve their effectiveness. To get to that point, we first provide in this section a brief review of general, domain-independent, system safety principles.

Detailed safety measures abound in each industry and for dealing with different hazards (e.g., electrocution, fire). In contrast with this proliferation of safety measures, there exist a small set of safety principles, which are domain-independent and technologically agnostic, and from which many safety measures derive. Said differently, a small set of general safety principles can be translated and adopted in many different ways as safety measures to deal with a broad range of hazards in different contexts. In this section, we briefly examine these high level safety principles. The distinction between specific safety measures and general safety principles is somewhat similar to that between *tactics* and *strategy* in a military context: the former relates to specific moves and dispositions to achieve a local objective (e.g., moving soldiers and equipment, engaging in a skirmish), whereas the latter, *strategy*, relates to broader considerations for planning and organizing to succeed in a general conflict (e.g., war) with an opponent. The following safety principles build on the notion of hazard level (and escalation) and accident sequence. They can be implemented in a variety of ways, and they require creativity and technical ingenuity to implement in different contexts and for handling different types of hazards. Although these safety principles are not meant to be exhaustive, they cover a broad range of safety considerations, and many detailed safety measures derive from or can be traced back to them.

### *3.1. The fail-safe principle*

The fail-safe principle requires or is defined by one particular solution to the problem of how a local failure affects the system hazard level. Consider for example the failure of a component in a system; this local event can propagate and affect the system in different ways. For example, it can lead to a cascading failure (domino effect), which would result in a complete system failure or accident (e.g., nodes in an electric power grids operating at maximum capacity). It can also remain confined to the neighborhood of the failed component and hence have a limited impact at the system hazard level. Specifically, the fail–safe principle requires design features such that the failure of a component in a system will result in operational conditions that (i) block an accident sequence from further advancing, and/or (ii) freeze the dynamics of hazard escalation in the system, thus preventing potential harm or damage. In this view, the fail-safe principle represents a particular form of robustness and failure tolerance. Conversely, if the fail-safe principle is not implemented, a component's failure would aggravate a situation by further escalating the system hazard level, thus initiating an accident sequence or leading to an accident (Fig. 7).

The fail-safe principle can be implemented in a variety of ways. For example, air brakes on trucks are maintained in the open position by pressure in the lines; should the pressure drop because of leakage or any other failure mechanism, the brakes will be applied. Another example of the implementation of the fail-safe principle is the "dead man's switch" for train operators: should they fall asleep or become unconscious, the device is no longer held down, and as a result the brakes are applied. More complex implementations of the fail-safe principle can be found in nuclear reactors where self-shutdown is initiated if critical operating conditions are reached. While there may be situations or items for which the

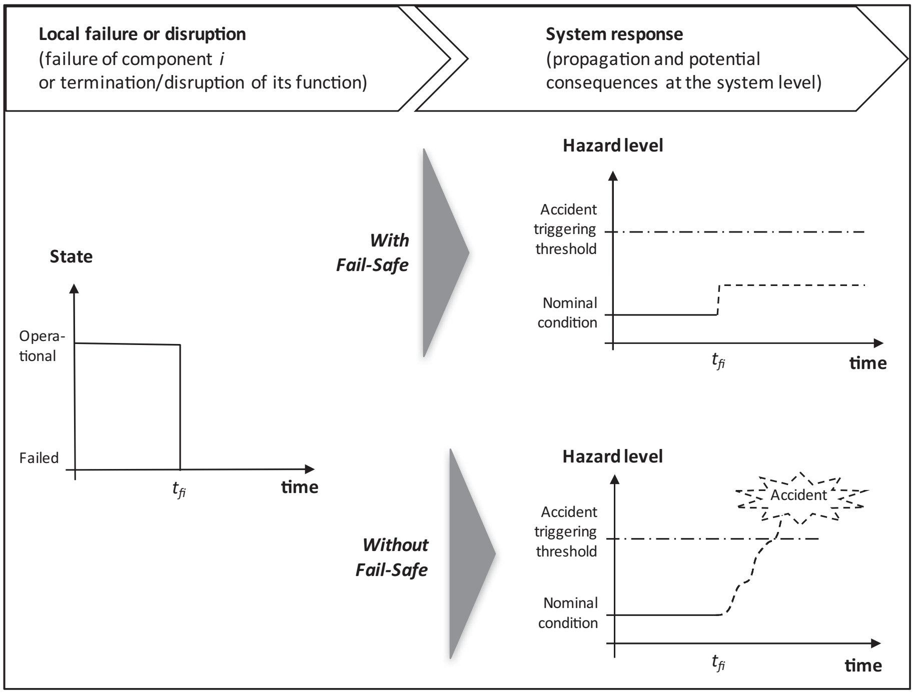


**Fig. 7.** Illustrative comparison of system hazard level following a local failure, both with the implementation of the fail-safe principle and without it ($t_{fi}$ is the time of occurrence of the failure of the component *i*).

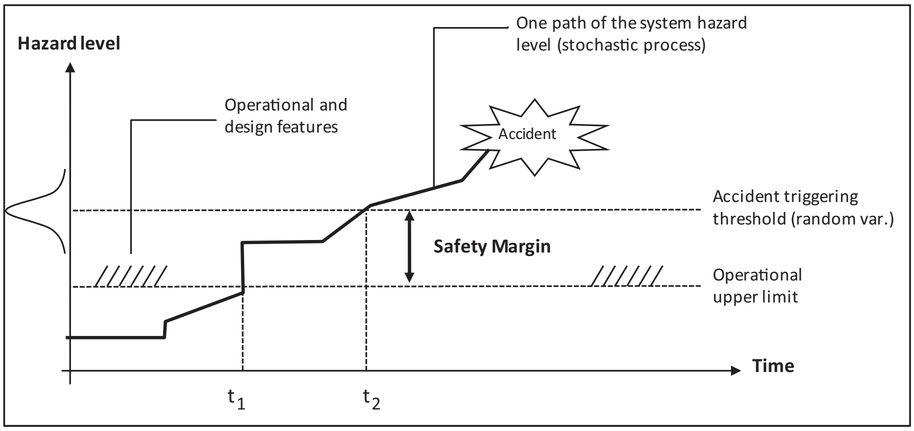


**Fig. 8.** Illustration of the safety margins principle with a sample accident trajectory from a nominal operating condition to an accident.

fail-safe principle is incompatible with their design or is simply not implementable, it is nevertheless important that this principle always be considered and carefully assessed in any design endeavor before it is ruled out.

### 3.2. The safety margin principle

The adoption of safety margins is a common practice in civil engineering where structures are designed with a safety factor to account for larger loads than what they are expected to sustain, or weaker structural strength than usual due to various uncertainties. The importance of safety margins for structures such as bridges and levees, which have to cope with the uncertainty of operational and environmental conditions such as wind force and wave height, is easy to understand. The idea of safety margins in civil engineering is an instantiation of a broader safety principle, which we refer to by the same name. The safety margin principle extends beyond civil engineering and is more diverse in its implementation than the particular form it takes for structures. It requires first an estimation of a critical hazard threshold for accident occurrence, and an understanding of the dynamics of hazard escalation in a particular situation. For example, methane in coalmines enters an "explosive range" when its concentration in the mine atmosphere reaches between 5% and 15% (Saleh and Cummings, 2011). Reaching the 5% threshold for example can be considered a critical hazard threshold in the mine**.** The safety margin principle requires that features be put in place to maintain the operational conditions and the associated hazard level at some "distance" away from the estimated critical hazard threshold or accident-triggering threshold (Fig. 8). For instance, in the coal mine example, a safety margin can be established with respect to the risk of methane explosion by maintaining methane concentration below say 3% in the mine atmosphere, 2 percentage points below the critical hazard level. The difference between the operational upper limit (3%) and the boundary of the explosive range (5%, the triggering threshold) is a particular form of safety margin in this context. **Safety margins are one way for coping with uncertainties in both the critical hazard threshold (a random variable) and in our ability to estimate and manage the actual operational conditions in a system**, such that their associated hazard level does not intersect with the real critical hazard threshold.[11]

### *3.3. The un-graduated response principle: rules of engagements with hazards*

The use of force in a military or law enforcement context is governed by a set of Rules of Engagements or Rules for the Use of Force (CJCSI, 2005). The principal tenet of these rules is that of a *graduated response*, namely that if force is deemed necessary, it ought to be applied gradually in relation to the extent of a demonstrated belligerence, and only the minimum force necessary to accomplish the mission should be used. The opposite of this tenet holds for dealing with safety issues, and the corresponding principle we refer to as the un-graduated response or rules of engagements with safety hazards. This principle for accident prevention and mitigation articulates a hierarchy of preferences for safety interventions. It posits that the first course of action to explore for accident prevention is the possibility of eliminating a hazard all together. We refer to this course of action as "kill first" or the use creativity and technical ingenuity as a first resort to eliminate the hazard, regardless of the extent of its *belligerence* (lethal use of force against hazards). For example, many precautions can be taken when transporting hazardous materials, such as the use of thicker and sturdier containers. But eliminating the hazard all together instead of better containing it, by transporting a safer substitute for example ought to be the first course of action to consider and examine for feasibility. Similarly, if a heat source or electric wires are in the vicinity of flammable material, the hazard can be controlled or the probability of an accident reduced by using proper wire isolation and placing the wires within fireproof protective jackets. But this particular hazard, the co-location of the electric wires and flammable material, can be eliminated by re-routing the wires through another location—the preferred course of action by virtue of this safety principle. Next, if the hazard cannot be eliminated, the second course of action is to control it or reduce its likelihood of escalating into an accident. Figuratively, if "kill first" is not feasible, then proceed to "apprehend and restrain". We revisit these issues in more details when we discuss the defense-in-depth safety principle next.

### *3.4. The defense-in-depth principle*

Defense-in-depth (DiD) is a fundamental safety principle and one whose importance cannot be underestimated. It derives from a long tradition in warfare by virtue of which important positions were protected by multiple lines of defenses (e.g., moat, outer wall, inner wall). First conceptualized in the U.S. nuclear industry, defense-in-depth became the basis for risk-informed decisions by the Nuclear Regulatory Commission (NRC, 2000; Sørensen et al., 1999-2000), and it is adopted under various names/forms in other industries. The principle has several pillars and requires that (i) multiple lines of defenses or safety barriers be placed along potential accident sequences; (ii) safety should not rely on a single defensive element (hence the "depth" qualifier in DiD); (iii) the successive barriers should be diverse in nature and include technical, operational, and organizational safety barriers. The various safety barriers have different objectives and perform different functions. The first set of barriers, or line of defense, is meant to prevent an accident sequence from initiating. Should this first line of defense fail in its prevention function, a second set of safety defenses should be in place to block the accident sequence from further escalating. Finally, should the first and second lines of defense fail, a third set of safety defenses should be in place to contain the accident and mitigate its consequences. This third line of defense is designed and put in place based on the assumption that the accident will occur, but its potential adverse consequences should be minimized. These three lines of defenses constitute defense-in-depth and its three functions, namely prevention, blocking further hazardous escalation, and containing the damage or mitigating the potential consequences (Fig. 9).

Accidents typically result from the absence, inadequacy, or breach of defenses. The notion of a safety barrier is the embodiment of the "defense" part of DiD in the sense that defenses are realized through barriers deliberately inserted along potential accident sequences and prior to their initiating events.

### *3.5. The observability-in-depth principle*

There are several mechanisms in the design and operation of complex systems that can contribute to the concealment of the occurrence of hazardous events or conditions (e.g., failures of redundant components or build-up of latent failures) and that the system has transitioned to an increasingly hazardous state. These mechanisms make "systems more [...] opaque to the people who manage and operate them" (Reason, 1997). As a result, system operators or users may be left blind to the possibility that hazard escalation is occurring, thus decreasing their situational awareness and shortening the time they have to intervene before an accident is released. Operators make decisions during system operation, which are based on and affect the hazard level in a system. If the system conditions/states are not carefully monitored and reliably reported, there is a distinct possibility that the hazard level *assumed* will diverge from the *actual* hazard level reached by the system. The gap between these two quantities can result in the operators making flawed decisions, which in turn can compromise the safe operation of the system or fail to check the escalation of an accident sequence (e.g., no action when an intervention is warranted, see for instance (Saleh et al., 2014b).

The observability-in-depth (OiD) safety principle is specifically meant to pre-empt these issues, and constitutes an important complement to defense-in-depth, without which the latter (DiD) can devolve into a defense-blind safety strategy. OiD requires and is characterized by the set of provisions, technical, operational, and organizational designed to enable the monitoring and identification of emerging hazardous conditions, accident pathogens, and adverse events in a system—to eliminate safety blind spots that might be introduced in the system because of DiD or other hazard concealing mechanisms. It requires that all safety-degrading events or states that safety barriers are meant to protect against be observable. This implies that various features be put in place to observe and monitor the system state and breaches of any safety barrier, and reliably provide this feedback to the proper stakeholders (operators, users, engineers, managers, etc.). OiD seeks to: (i) minimize the gap between the actual and the assumed hazard levels, and (ii) ensure that at the hazard levels associated with the breaching of any safety barrier, these two quantities coincide. The "depth" qualifier in OiD has both a causal and a temporal dimension, and it characterizes the ability to identify adverse states and conditions far upstream (early) in an accident sequence. It reflects the ability to observe emerging accident pathogens and latent failures before their effect becomes manifest on the system's output or behavior, or before a more hazardous transition occurs in an accident sequence. This principle does not affect or intervene

[11] Furthermore, a larger margin makes it more likely that the system state will not reach the accident-triggering threshold, or that a longer time window is available to detect a system state that has crossed the operational upper limit and abate the hazardous situation before an accident is triggered.

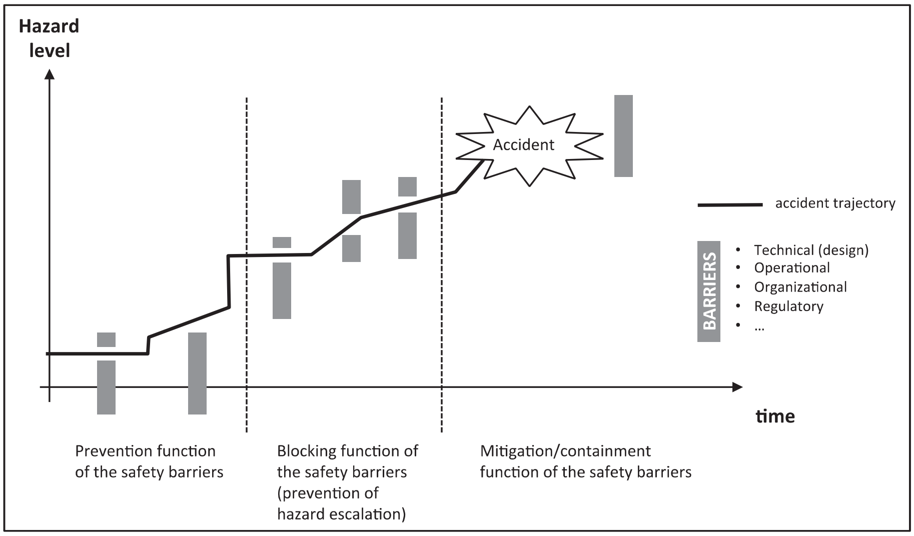

Hazard level
Accident
accident trajectory
BARRIERS
• Technical (design)
• Operational
• Organizational
• Regulatory
• …
time
Prevention function of the safety barriers
Blocking function of the safety barriers (prevention of hazard escalation)
Mitigation/containment function of the safety barriers

**Fig. 9.** Illustration of the defense-in-depth safety principle, along with a hypothetical accident sequence (its occurrence is the result of the absence, inadequacy, or breach of various safety barriers).

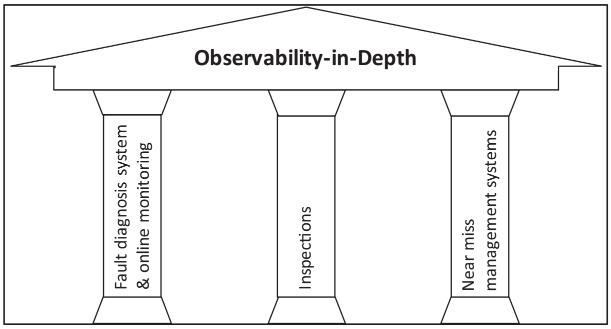

Observability-in-Depth
Fault diagnosis system & online monitoring
Inspections
Near miss management systems

**Fig. 10.** The three pillars of the implementation of observability-in-depth.

directly in an accident sequence, but it scans and monitors for hazardous conditions and escalation. Its significance is best understood by considering situations in which this principle is NOT implemented. Violations of the OiD principle highlight not the causal chain of an accident sequence (why the accident happened), but the causal factors that failed to support accident prevention (why safety interventions and the blocking the accident sequence did not take place).

In short, observability-in-depth is an information-centric safety principle, and it is an important ingredient in the development of a dynamic defense-in-depth safety strategy in which some defensive resources, safety barriers and safety interventions for example, are prioritized and triggered dynamically in response to identified emerging risks.[12] The parallels and synergies between observability-in-depth and near-miss management systems are evident and will be examined in the next section. We will examine the relevance and synergies with the other safety principles and how these can help inform the design and operation of near-miss management systems and improve their effectiveness.

## 4. Near-miss management systems, observability-in-depth, and safety principles

The safety principles discussed in the previous section offer a new lens by which to view near-miss management systems (NMS), and they can help inform and guide the two central phases in these systems, namely the *causal analysis* and the *solution identification* phases (Fig. 1). We will examine shortly the different modes by which the safety principles can be helpful in the design and operation of NMS, but before doing so, we discuss first the connections and similarities between observability-in-depth and NMS.

### 4.1. The three pillars of the implementation of observability-in-depth

OiD and NMS have significant overlap. Although they to operate on different levels of abstractions, it can be said that NMS is one of the pillars of the implementation of observability-in-depth (see Fig. 10). The two other pillars are "fault diagnosis systems and online monitoring", and regular "inspections."[13] These three approaches are information-centric and meant to scan for, detect, and assess adverse conditions and hazardous occurrences in a system before they escalate into full blown accidents. They operate though by different means and over different time scales as we discuss next.

Recall that OiD is characterized by the set of provisions, technical, operational, and organizational designed to monitor and identify emerging hazardous conditions, accident pathogens, and adverse events in a system. Its end-objective is to contribute to accident prevention by thwarting safety blind spots from settling in, and in so doing, it provides information about emerging hazards for dynamically prioritizing and allocating safety resources for targeted safety interventions. These features and objectives are information-centric as discussed previously, and they can be pursued and achieved in different ways:

(1) In real time by proper instrumentation, sensors, and algorithms to continuously (and autonomously) monitor the state of a system, and identify degradation and faults in components and equipment before they escalate into serious accidents. This is the purview of fault diagnosis systems.
(2) Asynchronously by qualified individuals (the "sensors") who target specific subsystems or equipment (and in some cases execution of operational procedures) to assess their conditions and suitability for continued use. This can subsequently trigger equipment maintenance or replacement for example. This is the purview of periodic inspection.

[12] Observability-in-depth echoes the old Russian say, "trust but verify". If the previous safety principles are meant to build (some) trust in the safety of a system, observability-in-depth is concerned with the "verify" part.

[13] Inspections can be calendar-based, clock-based, or condition-based.

(3) Asynchronously by anyone interacting with the system, operators, technicians, engineers, managers, etc. The "sensors" in this case are a host of individuals, not designated beforehand, and who happen to stumble on an adverse condition or experience an accident precursor or near-miss. The temporal onset in this case, unlike with inspections, is random or stochastic. This is the purview of NMS (the first couple of phases).

It is instructive to see how these three approaches support the broader umbrella[14] of observability-in-depth and constitute different implementations of its requirements. The differences between them are whether the monitoring is done in real time or asynchronously on the one hand, and whether the "sensors" are hardware/software or liveware on the other hand. Furthermore, while the "sensors" are narrowband and focused on pre-defined specific tasks with inspections, they are broadly distributed and *accidental* with near-miss management systems.

The boundaries between these three approaches (fault detection/online monitoring, inspection, and NMS) are likely to be blurred in the future, and the next generation of near-miss management systems, NMS 2.0, can usefully integrate them and fuse data from multiple sources to improve the efficacy of precursor identification, safety interventions, and ultimately accident prevention. For example, black boxes, sensors, and automation can become an integral part of near-miss management systems in the future. Vision-based sensors and safety equipment can also be integrated within these NMS 2.0, and machine learning techniques will help identify anomalous situations and excursions of operational parameters outside safety envelops. These technologies and some level of autonomy will complement the current approaches to NMS, which remain *manual* to some extent.

### *4.2. Violations of safety principle and near-miss management systems*

Prior to the devastating accident at the Texas City refinery in 2005 in which scores where killed and over 180 injured, several start-ups of the isomerization unit[15] where the accident occurred resulted in raffinate filling the tower above the first, and in some cases the second, recommended safety threshold. These instances constitute serious accident precursors, but they were not given any attention until the accident occurred (in roughly a similar manner for its initiating event). Had a proper NMS been in place and these precursors reported, they could have been examined in different ways and several lessons learned could have been extracted. We propose that these instances reflect fundamentally a violation of safety principles, in particular the safety margin principle, and to some extent the defense-in-depth principle.

We discussed previously the prioritization process of near-miss data and the importance of identifying the underlying generating mechanism of a given precursor or near-miss, not just its immediate cause or symptom, during the *causal analysis* phase of a NMS (2.2). The safety principles examined in Section 3 offer a new lens by which to view NMS, and they can help inform and guide the two central phases in these systems, namely the *causal analysis* and the *solution identification* phases. More specifically:

(1) All near-miss data should be sifted through the additional filter of compliance with or violation of safety principles (whether the high level principles here presented or specific ones tailored to the particular industry where the NMS is used).

(2) Repeated violation of a particular safety principle, for example the safety margin in the Texas City refinery accident, can be indicative of a fundamental precursor generating mechanism. The NASA accident precursor handbook emphasizes the importance of generalization during the causal analysis, "which helps to extrapolate lessons learned onto other systems or other scenarios... to go beyond the circumstantial aspects of the [precursor] as it occurred" (NASA, 2011). We propose that all precursor and near-miss data be examined in light of their potential violation of particular safety principles and be grouped or clustered together according to this feature.

(3) The clustering of precursors based on this new measure of similarity (violation of safety principle *j*) offers several advantages. First it can provide some indication regarding the safety culture at the plant or company. For example, a high number of violations of the safety margin principle, as was the case at the Texas City refinery, is likely to reflect both a poor safety culture and safety awareness, as well as deficient training. This in turn can help devise an appropriate safety intervention, or as noted earlier, identify the efficient way to kill the near-miss generating mechanism. The intervention can have positive repercussions across other instances of precursors within the same cluster.[16] Second, the clustering of precursors according to our proposed measure of similarity can help identify vulnerabilities in the design of the system. For example, repeated violations of the defense-in-depth principle can point out whether there are weakness or absences of safety barriers performing the prevention function (none existed at the Texas City refinery, or with the foam strikes on the space shuttle, yet they would have been easily implementable), or the blocking functions. Third, our proposed clustering scheme can help safety professionals and other stakeholders involved in the NMS at a company remain attuned to and vigilant with respect to violations of these safety principles (or an expanded set tailored to the particular industry). By doing so, they can identify best practices and help improve their handling of particular violations of safety principles.

(4) Examining precursors in light of violation of safety principles can also lead to actionable findings and guide the *solution identification* phase of a NMS. In this view, the objective of this phase becomes to (re-)establish and enforce violated safety principles by several means synergistically, for example through design solutions, improved operational procedures, and better training. By distributing the ownership of the safety principles and conformance with them across the enterprise, NMS and safety professionals can help improve the safety culture and awareness within a company, and by targeting the underlying precursor generating mechanisms, namely the violations of safety principles, they can help not only reduce near-misses, but ultimately improve accident prevention and sustain system safety.

## 5. Conclusions

This work brought together two strands in the literature on accident prevention and system safety: the examination of near-miss management systems on the one hand, and the development of general domain-independent system safety principles on the other hand. NMS are broadly adopted across different hazardous industries and increasingly so in other sectors such as manufacturing, construction, and healthcare. Their objective is to collect and

[14] The entablature and pediment in Fig. 10.

[15] Following shutdown for maintenance. Details can be found in Saleh et al. (2014b).

[16] The idiom, "to kill two birds with one stone" comes to mind, except in our case, it is the handling of *n* precursors within the same cluster targeted with one safety intervention.

prioritize anomaly and precursor data, to interpret and assess their risk implications, and to transform this data into risk-informed interventions and safety improvements and awareness. NMS are meant to tease out the failure generating mechanisms in, and the risk implications of, anomaly and precursor data and reflects them back to the organization in a variety of ways to have them addressed. Recognizing that learning from near-misses is less costly than learning from accidents, the main value of a NMS is in the learning loop it provides within and across organizations, in focusing safety resources on addressing unsafe acts, reducing unsafe conditions and procedures, and improving design and operational safety issues.

Since no common standard or guidelines exist for safety managers in this regard, a careful literature review and synthesis of key issues and challenges in designing and operating NMS was first provided in this work. Important findings were highlighted, for example:

- It was argued that NMS should not be confined to event-based safety incidents. Many safety blind spots would remain and several learning opportunities to improve safety would be forfeited by doing so.
- The case was also made that two different types of analyzes are important to consider during the *causal analysis* phase of a NMS: a retrospective analysis focused on identifying the causal basis of the precursor and its generating mechanism; and a prospective analysis of hypothetical futures extrapolating the precursors to accidents that could have happened.
- It was proposed that the military concept of a **kill chain** is useful to appreciate the key issues in this *solution identification* phase of a NMS: the objective of this phase should be to identify the most efficient way to kill the near-miss generating mechanism, to eliminate (to the extent possible) its key ingredients for all future times. Safety professionals are encouraged to avoid the temptation of quick *band-aid* solutions that do not address the fundamental generating mechanism of the near-miss.

This work then proposed and examined synergies between fundamental safety principles adopted in risk management, including defense- and observability- in depth, and NMS. Safety principles offer a new lens by which to view NMS. One important result is that near-miss data can be classified and interpreted in light of safety principles violated, and that safety interventions can be particularly effective when organized around such findings, the objectives being to (re-)establish and strengthen compliance with safety principles through workforce training, system redesign, and/or improved operational procedures. Finally, it was argued that NMS is one of the pillars of the implementation of observability-in-depth, and that the boundaries with the two other pillars (fault detection/online monitoring, and inspection) are likely to be blurred in the future, and that the next generation NMS (2.0) will likely integrate data from multiple sources to improve the efficacy of precursor identification, prioritization, and safety interventions, and ultimately accident prevention. Further developments will be oriented to validate the contributions of these safety principles through field testing.

## References

Bakolas, E., Saleh, J.H., 2011. Augmenting defense-in-depth with the concepts of observability and diagnosability from Control Theory and Discrete Event Systems. Reliab. Eng. Syst. Saf. 96 (1), 184–193.

Barach, P., Small, S.D., 2000. Reporting and preventing medical mishaps: lessons from non-medical near-miss reporting systems. Br. Med. J. 320, 759–763, 18 March.

Bird, F.E., Germain, G.L., 1996. Loss control management: Practical loss control leadership. Revised Edition. International Loss Control Institute, USA.

Cambraia, F.B., Saurin, T.A., Formoso, C.T., 2010. Identification, analysis, and dissemination of information on near-misses: a case study in the construction industry. Saf. Sci. 48, 91–99.

Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3121.01B, 2005. Standing Rules of Engagement/Standing Rules for the Use of Force for U.S. Forces.

Cowlagi, R.V., Saleh, J.H., 2013. Coordinability and consistency in accident causation and prevention: formal system-theoretic concepts for safety in multilevel systems. Risk Anal. 33 (3), 420–433.

Favaro, F.M., Saleh, J.H., 2014. Observability-in-Depth: an essential complement to the defense-in-depth safety strategy in the nuclear industry. Nucl. Eng. Technol. 46 (6), 803–816.

Goldenhar, L., Williams, L.J., Swanson, N.G., 2003. Modelling relationships between job stressors and injury and near-miss outcomes for construction labourers. Work & Stress 17 (3), 218–240.

Gnoni, M.G., Andriulo, S., Maggio, G., Nardone, P., 2013. "Lean occupational" safety: an application for a Near-miss Management System design. Saf. Sci. 53, 96–104.

Gnoni, M.G., Lettera, G., 2012. Near-miss management systems: a methodological comparison. J. Loss Prevent. Proc. Ind. 25 (3), 609–616.

Heinrich, H.W., 1931. Industrial accident prevention: A scientific approach. McGraw-Hill.

Hopkins, A., 2001. Was the Three Mile Island a "Normal Accident? J. Contingencies Crisis Manage. 9 (2), 65–72.

Hovden, J., Stroseth, F., Tinmannsvik, R.K., 2011. Multilevel learning from accidents – Case studies in transport. Saf. Sci. 49 (1), 98–105.

Jones, S., Kirchsteiger, C., Bjerke, W., 1999. The importance of near-miss reporting to further improve safety performance. J. Loss Prev. Process Ind. 12, 59–67.

Kleindorfer, P., Oktem, U.G., Pariyani, A., Seider, W.D., 2012. Assessment of catastrophe risk and potential losses in industry. Comput. Chem. Eng. 47, 85–96.

Manuele, F.A., 2011. Reviewing Heinrich dislodging two myths from the practice of safety. Prof. Saf., 52–61, October 2011

Masimore, L., 2007. Proving the Value of Safety Justification and ROI of Safety Programs and Machine Safety Investments. Rockwell Automation, USA. Available at <http://literature.rockwellautomation.com/idc/groups/literature/documents/wp/safety-wp004_-en-p.pdf> (accessed 10.10.11).

NASA, 2011. Accident precursor analysis handbook. NASA/SP-2011-3423. Washington DC. Available online at <http://www.hq.nasa.gov/office/ codeq/doctree/NASA_SP-2011-3423.pdf> (accessed 03.07.16).

National academy of Engineering, 2004. Accident Precursor Analysis and Management: Reducing Technological Risk Through Diligence. National Academies Press.

Nivolianitou, Z., Konstandinidou, M., Kiranoudis, C., Markatos, N., 2006. Development of a database for accidents and incidents in the Greek petrochemical industry. J. Loss Prev. Process Ind. 19 (6), 630–638.

Nuclear Regulatory Commission, 2008. Accident Sequence Precursor (ASP) Program Summary Description. 2008 Available online at <http://pbadupws.nrc.gov/docs/ML1319/ML13192A106.pdf> (accessed 25.05.16).

Nuclear Regulatory Commission, US, 2000. Causes and Significance of Design Basis Issues at US Nuclear Power Plants Draft Report. US Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington, DC.

Phimister, J.R., Oktem, U., Kleindorfer, P.R., Kunreuther, H., 2003. Near-miss incident management in the chemical process industry. Risk Anal. 23 (3), 445–459.

Reason, J.T., 1997. Managing the Risks of Organizational Accidents. Ashgate, Aldershot, Hants, England; Brookfield, Vt., USA.

Ritwik, U., 2002. Risk-based approach to near-miss. Hydrocarb. Process. 81 (10), 93–96.

Saleh, J.H., Saltmarsh, E.A., Favarò, F.M., Brevault, L., 2013. Accident precursors, near-misses, and warning signs: critical review and formal definitions within the framework of Discrete Event Systems. Reliab. Eng. Syst. Saf. 114, 148–154.

Saleh, J.H., Marais, K.B., Favaro, F.M., 2014a. System safety principles: a multidisciplinary engineering perspective. J. Loss Prevent. 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. Eng. Fail. Anal. 36 (2014), 121–133.

Saleh, J.H., Cummings, A.M., 2011. Safety in the mining industry and the unfinished legacy of mining accidents: safety levers and defense-in-depth for addressing mining hazards. Saf. Sci. 49 (6), 764–777.

Sattison, M.B., 2003. Nuclear accident precursor assessment: the Accident Precursor Program. In: Accident Precursor Analysis and Management. National Academy of Engineering, Washington DC.

Sørensen, J.N., Apostolakis, G.E., Kress, T.S., Powers, D.A., 1999. On the role of defense in depth in risk-informed regulation. In: Proceedings of the PSA '99. International Topical Meeting on Probabilistic Safety Assessment, Washington, DC, August 22–26, 1999, pp. 408–413.

Sterman, J.D., 1994. Learning in and about complex systems. Syst. Dyn. Rev. 10 (2–3), 291–330.

Van der Schaaf, T.W., Lucas, D.A., Hale, A.R. (Eds.), 1991. Near-Miss Reporting as a Safety Tool. Butterworth-Heinmann, Oxford.

Van der Schaaf, T.W., 1992. In: Moraal, J., Hale, A.R. (Eds.), Near-Miss Reporting in the Chemical Process Industry. Technische Universiteit Eindhoven, Proefschrift.

Van der Schaaf, T.W., Lucas, D.A., Hale, A.R. (Eds.), 2013. Near-Miss Reporting as a Safety Tool. Butterworth-Heinemann.

Van der Schaaf, T., Kanse, T., 2004. Biases in incident reporting databases: an empirical study in the chemical industry. Saf. Sci. 42 (1), 57–67.

Wu, W., Yang, H., Chew, D.A., Yang, S.H., Gibb, A.G., Li, Q., 2010. Towards an autonomous real-time tracking system of near-miss accidents on construction sites. Automat. Constr. 19 (2), 134–141.