LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TH-868930

IMPLEMENTATION OF OPERATIONAL AND COMPUTATIONAL SAFETY MODEL TO MITIGATE AND REDUCE INCIDENCE OF HIGHER SEVERITY EVENTS AT LAWRENCE LIVERMORE NATIONAL LABORATORY

R. Lara

September 2, 2024

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

IMPLEMENTATION OF OPERATIONAL AND COMPUTATIONAL SAFETY MODEL TO MITIGATE AND REDUCE INCIDENCE OF HIGHER SEVERITY EVENTS AT LAWRENCE LIVERMORE NATIONAL LABORATORY

by

Raul B. Lara Jr.

September 2024

Thesis Advisor: Matthew R. Crook
Co-Advisor: Steve Nakasaki, Lawrence Livermore National Laboratory

Distribution Statement A. Approved for public release: Distribution is unlimited.

Lawrence Livermore National Laboratories, the sponsoring agency for this research, approved release of this document (LLNL-TH-868930).

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2024	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
IMPLEMENTATION OF OPERATIONAL AND COMPUTATIONAL SAFETY MODEL TO MITIGATE AND REDUCE INCIDENCE OF HIGHER SEVERITY EVENTS AT LAWRENCE LIVERMORE NATIONAL LABORATORY

5. FUNDING NUMBERS

6. AUTHOR(S) Raul B. Lara Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Lawrence Livermore National Laboratory, Livermore, CA 94550

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
LLNL-TH-868930

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for public release: Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

High-severity events, such as serious injuries or fatalities, pose significant risks to workers, the environment, national security, and the reputation of national research laboratories, such as Lawrence Livermore National Laboratory (LLNL). A case study of 1,081 injury and illness cases was performed to assess the feasibility of implementing a framework to mitigate potential serious injuries or fatalities (pSIF) at LLNL. Review of the cases data quality showed that >86% of incidents had sufficient information to adopt the framework at LLNL.

Additionally, the case study reviewed institutional responses to the incidents. A computational model was developed to simulate pSIF incident distributions to deal with limitations from the case study, as well as to simulate institutional response. The findings concluded that while pSIF incidents were rare (<1% of total cases), the framework can improve organizational risk management by providing a consistent approach to incident response. It also suggests that resource allocation should focus on the highest risk areas, including noise exposure, overexertion, and repetitive motion. The computational model and framework offers a structured approach to reduce pSIF incidents, ultimately contributing to a safer research environment at LLNL. Although implementing the framework can enhance risk management, it requires commitment to quality data collection, incident classification, and integrated management systems for maximum efficacy.

14. SUBJECT TERMS
Lawrence Livermore National Laboratory, LLNL, safety, data modeling, Department of Energy, DOE, nuclear safety, health and safety, injury, illness, occupational health and safety, occupational health and safety management systems, OHSMS, industrial safety, nuclear, health, nuclear operations, research and development, potential serious injuries or fatalities, pSIF, fatalities

15. NUMBER OF PAGES
119

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Distribution Statement A. Approved for public release: Distribution is unlimited.

IMPLEMENTATION OF OPERATIONAL AND COMPUTATIONAL SAFETY MODEL TO MITIGATE AND REDUCE INCIDENCE OF HIGHER SEVERITY EVENTS AT LAWRENCE LIVERMORE NATIONAL LABORATORY

Raul B. Lara Jr.
Civilian,
BS, University of California, Santa Cruz, 2018
CERT OF COMPL, Air Force Institute of Technology, 2023

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN NUCLEAR OPERATIONS

from the

NAVAL POSTGRADUATE SCHOOL
September 2024

Approved by: Matthew R. Crook
Advisor

Steve Nakasaki
Co-Advisor

James H. Newman
Chair, Department of Provost

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

High-severity events, such as serious injuries or fatalities, pose significant risks to workers, the environment, national security, and the reputation of national research laboratories, such as Lawrence Livermore National Laboratory (LLNL). A case study of 1,081 injury and illness cases was performed to assess the feasibility of implementing a framework to mitigate potential serious injuries or fatalities (pSIF) at LLNL. Review of the cases data quality showed that >86% of incidents had sufficient information to adopt the framework at LLNL.

Additionally, the case study reviewed institutional responses to the incidents. A computational model was developed to simulate pSIF incident distributions to deal with limitations from the case study, as well as to simulate institutional response. The findings concluded that while pSIF incidents were rare (<1% of total cases), the framework can improve organizational risk management by providing a consistent approach to incident response. It also suggests that resource allocation should focus on the highest risk areas, including noise exposure, overexertion, and repetitive motion. The computational model and framework offers a structured approach to reduce pSIF incidents, ultimately contributing to a safer research environment at LLNL. Although implementing the framework can enhance risk management, it requires commitment to quality data collection, incident classification, and integrated management systems for maximum efficacy.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AIC	Akaike's Information Criterion
BIC	Bayesian Information Criterion
BFGS	Broyden-Fletcher-Goldfarb-Shanno
DOE	Department of Energy
EEI	Edison Electric Institute
EMS	Environmental Management System
ES&H	Environment Safety and Health
FMEA	Failure Mode and Effects Analysis
GLM	Generalized Linear Model
I&I	Injury and Illness
IAEA	International Atomic Energy Agency
LLF	Log-Likelihood Function
LLNL	Lawrence Livermore National Laboratory
NNSA	National Nuclear Security Administration
OHSMS	Occupational Health and Safety Management System
MEA	Mean Absolute Error
MLE	Maximum Likelihood Estimation
PAS	Preventative Action Score
PHLf	Final Potential Hurt Level
PHLwc	Potential Hurt Level Worst Case
PII	Personal Identifiable Information
PPE	Personal Protective Equipment
pSIF	Potential Serious Injury or Fatality
QMS	Quality Management System
SCL	Safety Classification Learning
SI	Sufficient Information
SIF	Serious Injury or Fatality

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to personally thank those that assisted in the case study that was conducted and would later become the basis of this thesis. Their support would not have made the case study or thesis possible. A portion of this thesis relied on a case study that was published and reused with permission from the co-authors. The author of this thesis was also the author and main contributor of the publication. Adapted with permission from *A Case Study in Assessing a Potential Severity Framework for Incidents from a Decadal Sample*. Copyright 2024 American Chemical Society.

Additionally, I would like to thank my loving wife, Ruth Yajaira Alvarez Mojica for her everlasting support. The foundation of research and science is not conducted without the support of countless hours from support staff, engineers, scientists, technicians, and in this case family. The staff at Lawrence Livermore National Laboratory also deserve praise with their continued support in research and development for national security and continuous improvement to ensure effective stewardship of government funds.

This work was performed under the auspices of the U.S. Department of Energy (DOE) by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

THIS PAGE INTENTIONALLY LEFT BLANK

DISCLOSURE STATEMENT

The thesis advisors and I agreed to limited use of generative artificial intelligence (AI) to aid in the process of improving grammar, punctuation, and concision. A large language model based on Open AI's GPT-4o mini was used to find and correct grammar, generate LaTeX code for some of the equations, assist in trouble shooting Python code and increase clarity in the writing. The paragraphs were used as the prompt and recommendations were given by the model or snippets of code with the error were given asking for suggestions in troubleshooting. To reduce the risk that the model introduced language that was false or derived from other works, each section or suggestion was reviewed and revised. Code was evaluated to ensure it worked as intended throughout creation and was edited to meet the objective of the program. Drafts were discussed with the Naval Postgraduate School Graduate Writing Center coach and approaches were discussed with colleagues to minimize the risk of writing or code that did not enhance or benefit the thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The Department of Energy (DOE) plays a critical role in ensuring security of the United States by addressing “energy, environmental, and nuclear challenges through science and technology solutions” [1]. Work of this nature involves cutting edge science and experimentation with often highly hazardous or catastrophic consequences if not performed safely. Given the proximity of LLNL to the community of Livermore, California, these consequences are magnified in the minds of residents and local officials out of concern for public safety and environmental protection. Consequently, it is imperative for LLNL to operate with utmost safety and aggressively implement controls to ensure workers are protected and public confidence is maintained. This thesis explores the efforts at Lawrence Livermore National Laboratory (LLNL) to mitigate high-severity incidents that pose risks to worker safety, the environment, and national security.

The focus of this thesis is on an operational framework designed to categorize the potential severity of occupational injuries and illnesses following incidents. Chapter I outlines the parameters of a case study that was conducted and highlights key occupational health and safety theories that have shaped safety management systems.

Subsequently, Chapter II details the methods and research associated with the case study. An evaluation of the data quality regarding incidents was conducted to determine the feasibility of a model focused on potential serious injuries or fatalities (pSIF) at LLNL. Chapter III discusses any noteworthy results from the case study as well as the limitations relevant to the creation of a computational model of the framework.

Chapter IV provides an overview of a computational model developed to simulate incident distributions from the case study, addressing limitations of the representative sample to better understand the pSIF model and its operational impact. Additionally, the institution’s responses to the incidents were modeled to identify opportunities for improvement in the case of framework implementation. The overarching framework proposed implements pSIF classification of incidents and considers their potential severity to inform institutional response based on those classifications.

Finally, Chapter V concludes with remarks on framework implementation and further research needed to address limitations of both the case study and the overarching framework. Understanding these constraints through the lens of the case study's context is crucial for interpreting relevant findings. Figure 1 illustrates how evaluation of the pSIF model feasibility at LLNL, evaluation of incident response, and the opportunity for performance improvement are used to inform framework evaluation at LLNL.

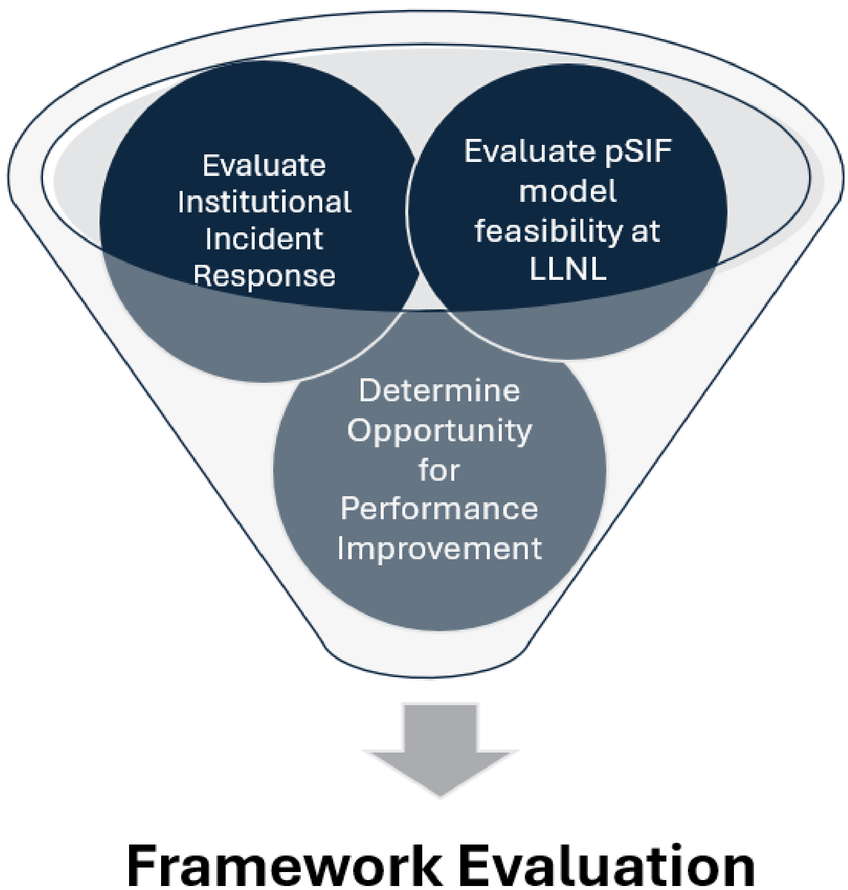

Figure 1. Steps Required for Framework Evaluation

A. IMPORTANCE OF INCIDENT MITIGATION

Public perception could quickly change in the event of an accident that permanently maims a worker. One notable incident involves Cecil Kelly, a chemical operator who was irradiated after a tank containing plutonium-239 dissolved in a chemical reagent. This lead to the release of nuclear energy, irradiating the operator and resulting in death within 35 hours of the incident [3]. Incidents involving criticality or other nuclear operations would not be tolerated by the public, especially if it were found to be completely preventable. The framework proposed is aimed to mitigate pSIF incidents through post-incident evaluation of hazards, controls, and likelihood of less severe incidents. Understanding the mission of LLNL allows a better understanding of why this framework is both useful and challenging to implement.

The mission of LLNL is broad and requires a multidisciplinary approach to its science and technology development, which involves taking on risk. As such, LLNL has core competencies in various fields of studies such as chemical, explosives, laser, nuclear, and emerging technologies [5]. Each core competency introduces a spectrum of hazards that necessitates a commitment from the laboratory to protect the environment, workers, and the public, requiring an integrated safety management system capable of identifying hazards and risks. Additionally, it entails various other management systems to aid in the completion of work, such as quality, environmental, and security.

A lapse in any of these areas could lead to a high severity incident affecting the health and safety of workers and the surrounding community. Such incidents can cause irreparable harm to individuals and their well-being, which in turn can undermine public trust in the responsibilities of the DOE and other national laboratories. This challenge is not unique; the DOE, like many organizations, must balance its operational needs with ensuring safe conduct of work.

Private industry and government agencies have a vested interest in not only mitigating adverse incidents, but also preventing pSIF incidents from occurring altogether. A risk management model can be particularly useful in this context. For example, the Navy regularly performs detailed analyses of flight mishaps, and the Naval Postgraduate School has conducted thesis work on correlation analysis of aviation mishaps as recently as 2023 to track and trend precursors or indicators [6]. Industries across the U.S. have invested tremendous resources in mitigating and preventing operational mishaps. The FAA has touted a 95% decrease in commercial aviation fatalities using a comprehensive and risk-based safety oversight process involving many facets of safety management and compliance [7]. Therefore, it is imperative that LLNL implement a framework focusing on pSIF to supplement its current methods. By doing so, LLNL may prevent pSIF incidents from occurring and continuously improve the performance of its safety management system.

B. BACKGROUND ON HEINRICH'S SAFETY TRIANGLE

There are various analytical approaches that have been proposed over the years, but one that still holds value in the health and safety community is the work that was done by Herbert William Heinrich in the 1930s. The foundational concept now known as Heinrich's Safety Triangle was quintessential in relating the occurrence of serious injuries and fatalities (SIF) proportional to a larger number of minor accidents, near misses, and unsafe acts [8]. Figure 2 illustrates a simplified Heinrich safety triangle with three categories, and although it was originally developed to relate the number of incidents at various severity levels, these distributions cannot be generalized to a particular industry and are often adjusted by industry [9], [10]. For this reason, the model is now largely used to describe the phenomenon that there is generally a larger amount of lower severity incidents and a much smaller amount of SIF incidents with no set ratio between severity. This theory still forms the conceptual basis of many organizations' safety management systems, although each level of the triangle may be interpreted differently.

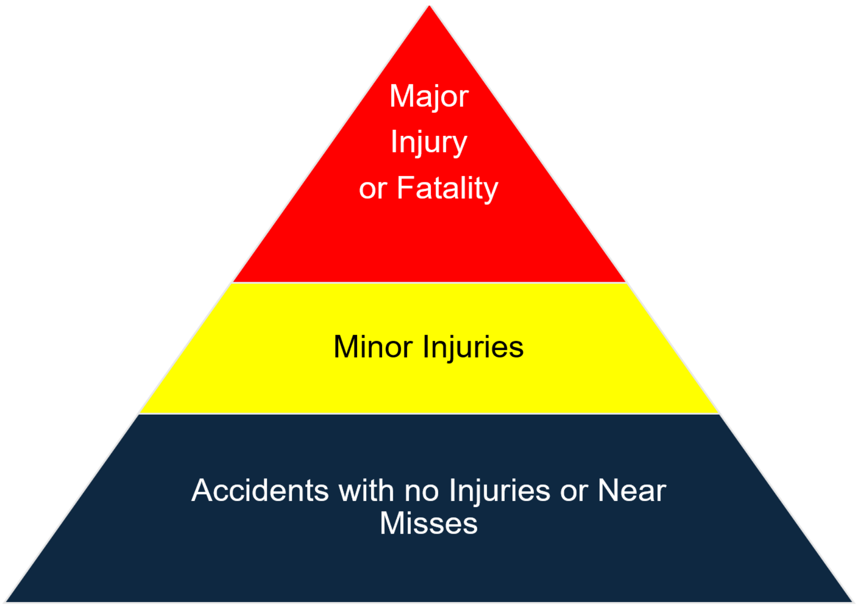

Figure 2. Heinrich's Safety Triangle Displaying Three Major Categories. Adapted from [8].

There has been a general decline in workplace injuries and illnesses over the years as efforts have been made to reduce unsafe acts, near misses, and minor injuries [11]. However, the same cannot be said about the rate of occupational fatalities illustrated in Figure 3 [11]. This is contrary to the Heinrich principle, suggesting that other factors or variables may be causing the SIFs to occur. For this reason, improvements in safety management systems based on this theory are necessary to better understand those other factors or variables in order to achieve the objective of preventing SIF incidents from occurring.

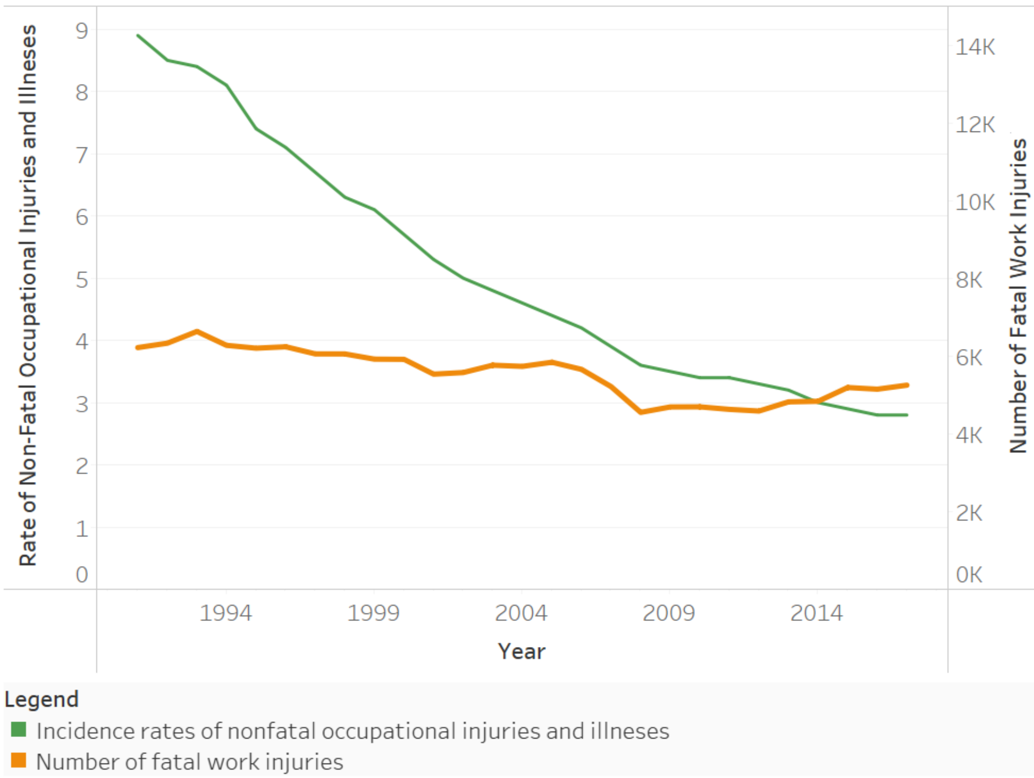

Figure 3. Occupational Injury and Illness Rates and Fatal Work Injuries. Adapted from [11].

C. PSIF MODELS IN PRIVATE INDUSTRY

Recently, there has been interest in private industry entities looking to mitigate SIF incidents. The Cambell Institute found that incidents had unique precursors to determine an incident's potential to become a SIF incident [4]. Additionally, the importance of implementing a method of reporting and analysis has been discussed as a topic of laboratory safety reforms in journals [12]. The case study highlighted in this thesis leveraged aspects of several models and is the majority focus of this work.

An important distinction between a model that looks at potential severity of events and Heinrich's Safety Triangle is that not all near misses or unsafe acts have equal correlation to a SIF. This distinction highlights the need to differentiate between events based on a pSIF classification model. However, this pSIF model must be accompanied by an adequate institutional response to create an effective framework.

Resources must be adequately allocated, as implementation of high-cost mitigation strategies may not reduce the number of SIF events. The importance extends beyond chemical laboratories and extends to the oil and gas industry. ExxonMobil's "Mining the Diamond" initiative examines the well-known safety triangle to focus on critical events based on its potential as opposed to the actual hurt level of the incident [13], [14]. To account for potential outcome of an incident, the safety management system must incorporate additional information for tracking and trending purposes. Figure 4 illustrates the shift in focus from medical treatment to identifying precursors that could result in a pSIF. Those incidents with SIF precursors should then be the focus of any institutional response.

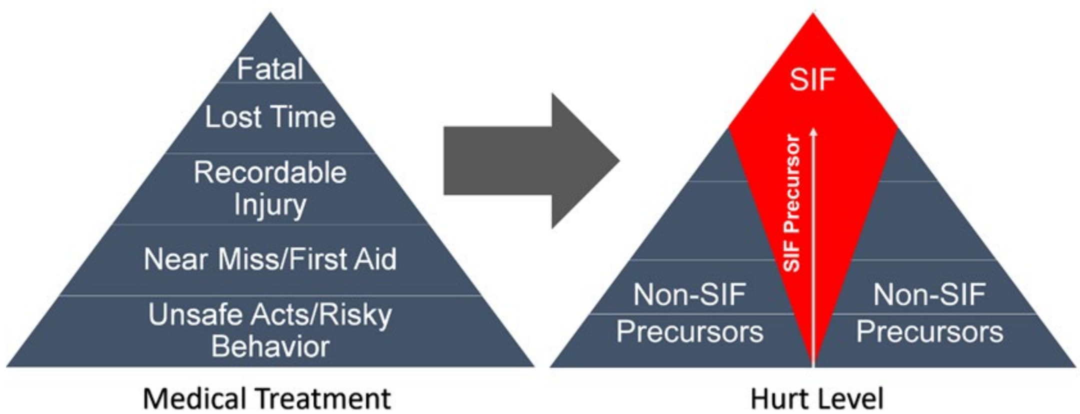

Figure 4. Shift from the Heinrich Safety Triangle to a Potential Severity Model. Source: [15].

The Edison Electric Institute (EEI) Safety Classification and Learning (SCL) model is used as a tool to define safety incidents to redirect attention from lower-severity events to those that could have been life threatening [16]. The EEI SCL model uses a decision-tree based approach to define and categorize safety incidents being predicated on the presence of high energy, application of controls, and the incident's outcome.

A 2021 white paper from DEKRA also described two approaches for defining an incident's potential severity. The first approach is a judgement-based narrative review which relies on safety practitioners to identify and analyze incidents. The second approach is an event-based decision tree classification approach that recognizes some activities produce higher proportions of precursor events [17]. Another study [18] investigated major accidents to find the most frequent contributing causes but was also industry-specific. Because various case studies and frameworks rely on industry-specific information, no definitive model is used across industries and firms for a pSIF classification tool. However, each model relied on several data elements from the incidents to make the determination [14], [16], [17], [18]. These elements included causal information, environmental factors, hazards present, and safety controls. For this reason, a custom model was developed for Lawrence Livermore National Laboratory that can be implemented and refined while aligning with the organization's collective philosophy.

D. EXISTING PROCESS AT LLNL

Currently at LLNL, there is capability to identify and address incidents through the occupational injury and illness (I&I) program as well as the institutional Near Miss program. However, there is currently no system that is specifically designed to address pSIF identification, tracking, and mitigation. The I&I program works in conjunction with an on-site clinic which includes both work-related and non-work-related cases. This will be relevant in evaluating the separation between occupational hazards and those related to recreational activities. The program investigates each incident to understand root causes, contributing factors, and determine corrective actions. Generally, corrective actions are implemented to prevent similar incidents in the future. During the investigation, if an incident uncovers issues, the investigator may prioritize corrective actions, including escalation to the organization's assurance manager for inclusion in an internal issue tracking system. This provides a documented institution's response to an incident. However, there are no standardized criteria for pSIF classification of an incident, and the evaluation relies heavily on the medical treatment resulting from the incident.

Near misses follow a separate workflow. Academia and other institutions often have near miss reporting systems that leverage self-reported incidents that did not result in an injury and are thought to be an important aspect of any safety management system [19]. According to LLNL's event notification and reporting document, workers involved in or witness a near miss are required to report the incident to their supervisor. If there is continuing potential for more serious consequences from the condition causing the near miss, the supervisor must contact the Environment Safety and Health (ES&H) Team, the Traffic Safety programs, facility owners or point of contacts, or the Infrastructure and Operations directorate to mitigate and correct the hazards. Conditions causing near misses and necessary corrective actions may escalate into the institution's issue tracking system.

Furthermore, the near miss may also meet the occurrence reporting threshold described in "Reporting Occurrences to DOE," where the Livermore Field Office—under the U.S. Department of Energy, Office of Inspector General—would oversee a contract response to severe incidents and provide concrete opportunities to reform the safety and health programs [20]. The involvement of government oversight offices to assist in government-owned, contract-operated institutions further highlight the importance of identifying those precursors that would cause the most harm to the environment, workforce, and government property. Additionally, government oversight may provide insights into opportunities for improving other management systems.

THIS PAGE INTENTIONALLY LEFT BLANK

II. METHODS AND RESEARCH

The following chapter has sections and information that is adopted from previously published in the journal American Chemical Society (ACS) Chemical Health and Safety [15]. Lawrence Livermore National Laboratory (LLNL) has an Injury and Illness (I&I) program that collects information on adverse incidents to ensure timely reporting and notification to federal oversight agencies, including the Department of Energy (DOE) Office of Environment, Health, Safety, and Security Occurrence Reporting and Processing System (ORPS) [24]. For the case study, an incident is defined as an event resulting in injury, illness, property loss, or environmental damage. The recorded data elements of the incident included those required by DOE form 5484.3 and the Department of Energy Injury and Illness Reporting Guide [25], such as the nature of the incident, description of the activity, corrective actions, and other relevant information. Additionally, LLNL information systems must comply with 10 C.F.R. 851 and the reporting requirements in Title 29 C.F.R. Part 1960 Subpart I, as well as DOE Order 231.1B Environment, Safety, and Health Reporting [26]. Adhering to these statutes and regulations ensures that there is uniform reporting for incidents at LLNL, facilitating analysis and testing of various safety models and frameworks with a standardized set of data. Figure 5 shows a generalized process for incidents where the data elements are recorded, and corrective actions are based on the severity of the incident.

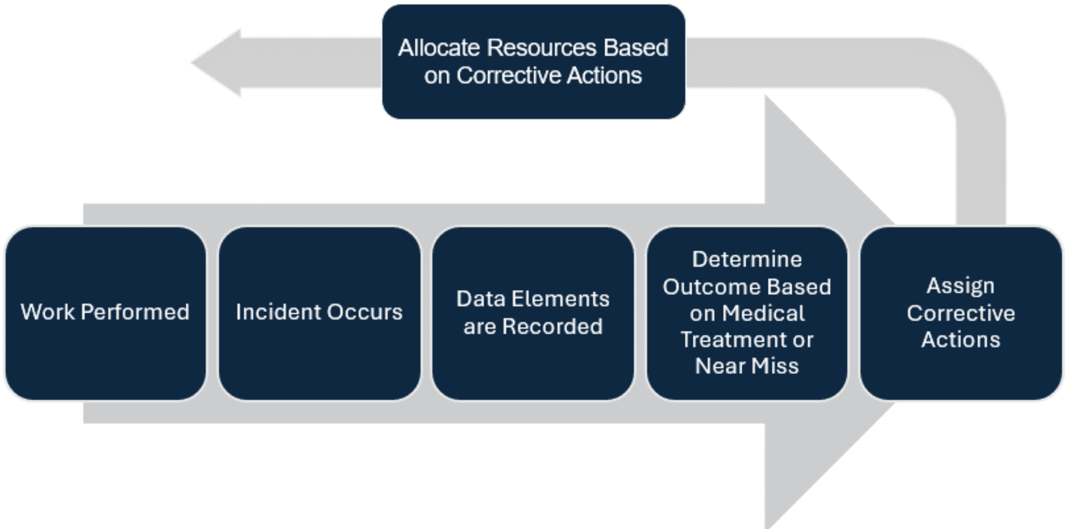

Figure 5. Current Generalized Process for Incidents

A. DATA SET TO TEST THE PSIF FRAMEWORK

An important focus for the case study was to assess the feasibility of adapting a pSIF framework to the current process at LLNL. Minimizing the impact to the current incident process would increase the feasibility of adopting the framework to the current process. Additionally, going through a sample of case studies through the proposed framework may highlight potential deficiencies in current process methodology. Figure 6 shows the proposed framework, with the changes in methodology highlighted compared to the current process seen in Figure 5.

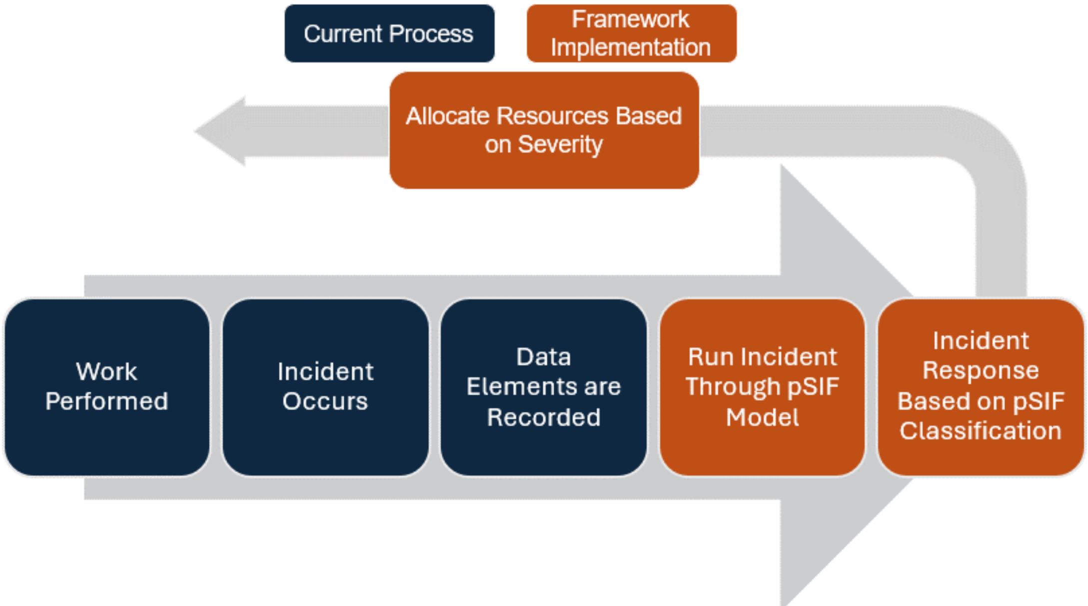

Figure 6. Proposed Change in Process for Incidents

The case study focused on I&I data from the calendar years 2007 to 2022. The case study attempted to look at incidents from prior to 2007, but only data from 2007 and on were found to be formatted in the same manner. Having a consistent dataset was crucial to understand whether a new process could be adopted to the current process and determine its feasibility. Between 2007 to 2022, approximately 5,700 incidents were recorded to the I&I system. Approximately 20% of the incidents were determined to be non-work related, as they occurred during recreational activities both on and off-site and during non-working hours.

With non-work-related incidents excluded, there was a sample size of 4,900 incidents [15]. This was done to ensure the framework was evaluated against occupational related incidents and to prevent unintentionally biasing the data by including hazards related to recreational activities. Because the focus is to mitigate occupational hazards, it seemed inappropriate to include recreational hazards that are accepted by the public.

The original dataset included a data element to describe the accident type, which comprised of 60 different types of accidents. However, many of these categories overlapped in terms of identified hazards. This is important in order to identify precursors associated with incidents. For example, one of the largest categories was found to be overexertion. It could be categorized as "overexertion by pulling or pushing objects," or "overexertion by lifting" [15]. But grouping the incidents also helped understand the types of hazards that are associated with injuries. The 60 accident types were further grouped by their categories where the population of those incidents was less than one percent of the total. However, incidents where there may be under sampling or unique hazards, such as radiation, were kept as their own category.

Due to the impracticality of running the model through each individual incident, a random representative sample of the incidents were selected for evaluation. To create a representative sample for the case study, the 60 different accident types for each incident were qualitatively grouped into 21 accident categories. Although the categories were grouped by hazard similarities or population size, not all categories could be grouped due to significant differences in hazard identification. Nevertheless, the differences between the grouped accident types were believed to accurately represent the total sample population.

Because the framework relies on looking at all incidents, regardless of medical treatment, there was no attempt to sample based on medical treatment. For example, an incident with the accident type of vehicular accident that leads to days away from work was weighted the same as a vehicular accident that lead to restricted workdays.

It is useful to understand the distribution of certain types of accidents and how it compares to other national laboratories. In doing so, one can evaluate if there are particular hazards or that particularly effect one site over another. For this reason, it is important to note that the distribution of accident types was not significantly different from other national research laboratories, such as the Los Alamos National Laboratory Plutonium Facility (TA-55) [27]. The distribution of incidents also did not differ from the DOE yearly Operating Experience Summary [28]. To avoid bias towards accident types that are more prevalent in the sample set, the accident type was used to distinguish the variance in the dataset. Accident types were also used because they may encompass incidents with similar hazards applicable to LLNL. To determine the sample needed from each category, the population variance of each grouped accident type was calculated using Equation 1 [29]. A 95% confidence level with a 5% margin of error was used to balance the precision and practicality of analyzing the sample data.

$$n = Z^2 \frac{p(1-p)}{ME^2} \tag{1}$$

Where:

- n is the number of incidents needed in the random sample to represent the entire population.
- Z is the standard score with a value of 1.96 for a confidence level of 95% to convert the score in a normal distribution to a standard normal distribution.
- ME is the Margin of Error to describe the amount of random sampling error in a survey, in this case five percent.
- P is related to the proportion of a category to the sample set.

Because the population size is finite, an additional correction factor must be applied as seen in Equation 2 [29].

$$n_{cor} = \frac{n}{1 + \frac{n}{N}} \tag{2}$$

Where:

- n is the number of incidents from Equation 1
- N is the total population size of the whole sample set

Equation 1 and Equation 2 were then used in a Python script to calculate the variance of the finite population. The results are seen in Table 1, where each incident was put in a list and N number of incidents were chosen at random. The list of N number of incidents for each accident category is also shown in Table 1 and is the random sample population for the case study.

B. MEASURING QUALITY OF RANDOM SAMPLE SET

With the data set sample chosen for the case study, the data quality of each incident needed to be assessed. Therefore, a method to measure the quality was developed to ensure that the sample data was not only representative, but sufficient to use in the case study. This would also highlight any opportunities for improving in the current process at LLNL for framework implementation.

Table 1. Incident Sample Size by Grouped Accident Types. Source: [15].

Grouped Accident Category	Total Incidents	Random Sample
Overexertion	684	145
Fall (all types)	657	140
Repetitive Motion	634	135
Struck by Object	479	106
Ingestion of substance	459	101
Rub or Abrasion	273	62
Contact	266	60
Noise Exposure	252	57
Bodily Reaction	250	57
Not Specific	190	43
Assault or Injury by Animal	185	42
Vehicular Accident	179	41
Caught in Equipment, materials, or machinery	114	26
Inhalation of substance	89	20
Walking/Running	64	15
Self-inflicted injury	28	6
Environmental or Object Temperature	27	6
Stationary Injury	26	6
Stepped on object	24	6
Radiation	20	4
Other*	13	3
Total	4,900	1,081
***Other includes the following grouped accident types: Reaction when surprised, frightened, startled; Air pressure changes exposure; Explosion; Fire -- unintended or uncontrolled; Welding Light Exposure**		

The variation in qualitative data would provide the context necessary in assessing the framework, a method was created to evaluate the qualitative data. The manual processing of qualitative data analysis and aggregates has been used in several studies with data collection methods including questionnaire data [30]. By measuring the quality of the random sample set a determination could be made if the current injury and illness program had sufficient information to run an incident through the pSIF model as well as evaluate the adequacy of the sample size. Each incident was assessed and given what will be noted as a sufficient information (SI) score, derived from four distinct criteria. Each criterion was assessed using a Boolean value with the intent to simply assess whether the information was present and valuable in the data elements. This provided a way to aggregate individual scores to allow for a quantitative assessment of data quality to facilitate comparisons between criteria. A data element was considered insufficient if the pSIF model practitioner must make assumptions or speculate without objective evidence.

The four distinct criteria were as follows and adapted from [15]:

- **Environmental Factors (EF)**: Details regarding the location, time of occurrence, and work environment.
- **Hazards Identification (HI)**: Details concerning the hazards present at time of incident.
- **Controls in Place (CIP)**: Identification and details of measures implemented, to include engineering controls, administrative protocols, personnel protective equipment, etc.
- **Causal Information (CI)**: Details pertaining to the factors contributing to the incident's occurrence, including both direct and indirect causes.

Each criterion was evaluated separately, and the summation of the individual scores became the SI score as seen in Equation 3.

$$SI_{incident} = \sum (EF_{incident} + HI_{inc} + CIP_{incident} + CI_{incident}) \qquad (3)$$

Data elements for the incidents, such as accident type, part of body, nature of injury or illness, at risk behavior, description of the activity, sequence of events, causal factors, immediate actions taken, corrective actions, and applicability of the integrated safety management system [31] – were all evaluated to determine whether the incident had sufficient information to inform each criterion. The SI score was used to determine a qualitative confidence level for each incident. The higher the score, the more likely a practitioner was able to determine whether the incident was a pSIF.

Table 2. SI Score and Confidence Levels

SI Score	Confidence Level
4	The practitioner has a high confidence in making a pSIF determination.
3	The practitioner has a medium confidence in making a pSIF determination.
<2	The practitioner has a low confidence in making a pSIF determination

The incidents that had an SI score of less than two resulted in a low confidence to no confidence determination of information sufficiency for the incident to be used in the pSIF model. If the majority of the incidents had key data elements missing, it would prevent meaningful analysis of relevant precursors. However, the statistics of the sample population, summarized in Table 3, showed that of the incidents from the sample, approximately 87%, had an SI score of 4, meaning that sufficient information was available for all four criteria.

Table 3. SI Score Statistics Summary. Source: [15].

Sufficient Information Score Sum	Number of Incidents	Relative Percentage to Sample
0 – 2	48	4.4%
3	97	9.0%
4	936	86.6%
Total	1081	100%

Because the aim of the framework is to improve the ability to identify opportunities to minimize potential serious injuries or fatalities, it was crucial to analyze the distribution of the SI categories to highlight any potential gaps in the data collection process. The total of the score for each criterion across the sample set determines the percentage of incidents that had met the criteria for each of categories. Table 4 shows the breakdown with Hazards Identification and Causal Information being within a tenth of a percentage point of each other. Although no category scored significantly lower than the others, and because greater than 95% of the incidents had an SI score of three or greater, the representative sample is determined to be adequate to test the framework.

Table 4. SI Score Category Summary. Source: [15].

Sufficient Information Category	Sum of Score	Relative Percentage to Sample Size
Hazards Identification	995	92.0%
Causal Information	996	92.1%
Controls in Place	1040	96.2%
Environmental Factors	1062	98.2%

C. THE OPERATIONAL PSIF MODEL

An operational pSIF model was developed as a tool in part of a broader framework to identify pSIF incidents and inform institutional response. For the purpose of the model, a pSIF score was derived from a 5-step process as illustrated in Figure 8.

1. Determine the potential worst-case outcome of an incident.

2. Determine the primary variable that would cause the potential worst-case outcome, considering the likelihood.

3. Determine the effectiveness of controls during the incident.

4. Use both the primary variable and effectiveness of controls to determine the V/C score.

5. Final pSIF determination, noted by PHLf.

The primary variable was the variable that could have resulted if the presence of the single variable were there and the likelihood of that variable occurring in light of mitigative controls in place. Review of incidents found several instances where more than one variable could have contributed to the reasonable worst-case outcome. Only one variable was chosen as the single-most valuable variable that would have contributed to the adverse effects of the event; this is recognized as a limitation and restriction of the model as often several variables contribute to an adverse event. A job-aid was developed for pSIF determination and is illustrated in Figure 7.

In order to simplify the model for use by practitioners in the field, a decision tree was created. Although one primary practitioner reviewed the majority of the incidents, a total of three practitioners reviewed incidents with higher pSIF scores in the sample dataset using the decision tree of Figure 7, as well as the data elements described in Section 2.B. Each of the practitioners held institutional knowledge of the specific work and hazards relating to the work they support. However, they were advised to make no assumptions of information from the incident that could not be verified using objective evidence. The following sections will describe each step of the decision tree. An example scenario of going through the decision tree is provided in Appendix A.

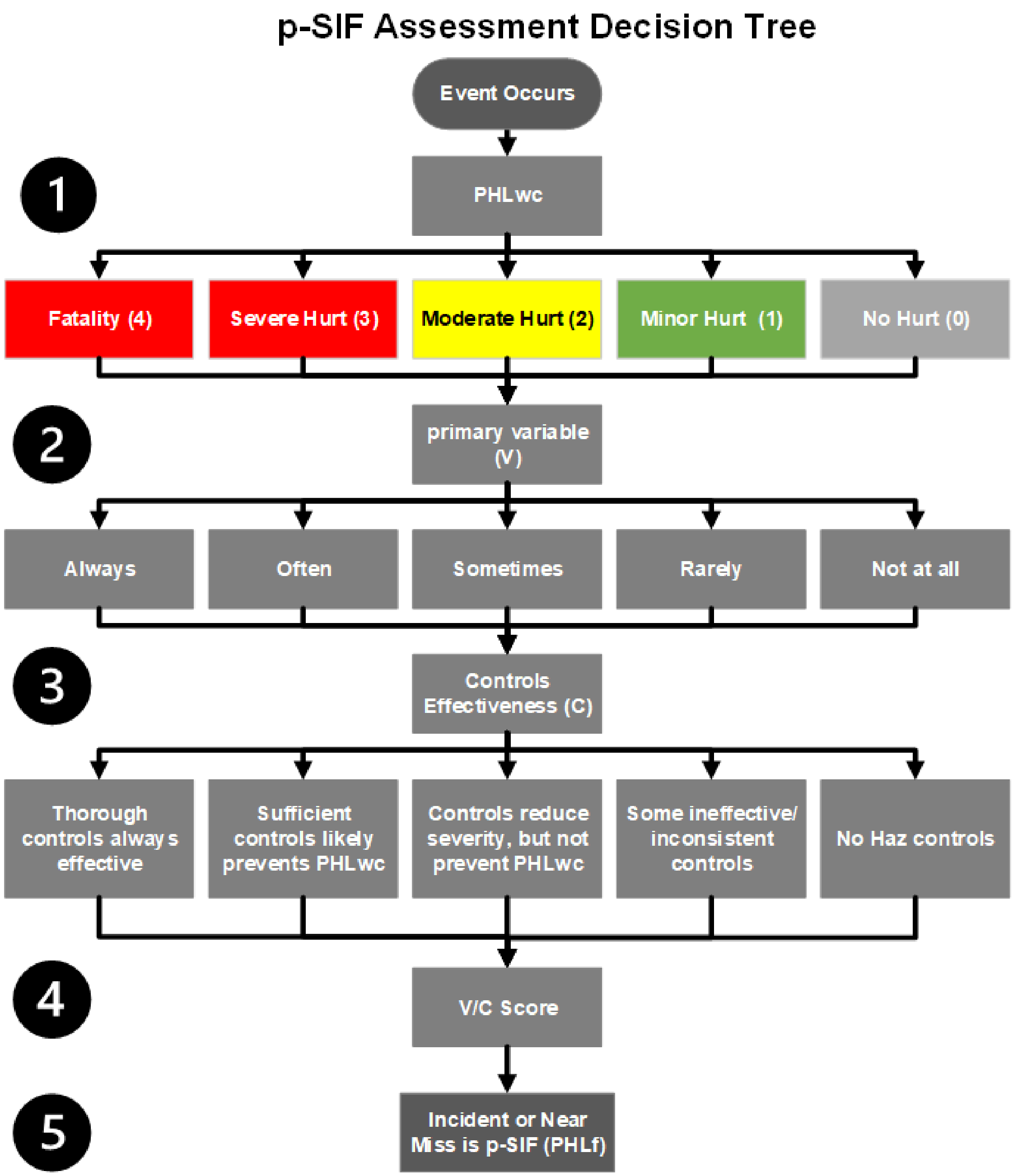

Figure 7. Decision-Tree Job Aid. Source: [15].

D. DETERMING HURT LEVEL

It is important to understand that time away from work after an incident is not indicative of the severity of an incident. For example, a person working in a laboratory setting receives chemical exposure leading to a chronic disease with life-long implications, while another person slips and breaks their arm. In these two scenarios, it is possible that both individuals may have similar time away from work and both produce a recordable injury, but one has clear long-lasting implications. This highlights a deficiency in focusing purely on medical treatment as opposed to the actual harm experienced by individuals in incidents.

Clear definitions of the level of harm and long-term health implications are needed to move away from a medical treatment focus. For this reason, the framework required a definition of hurt severity, along with examples. This was adopted from the EEI model that also gives examples of harm severity [16]. These definitions and examples are provided in Table 5 and were used to determine the harm level of a worst-case scenario.

Practitioners followed the framework outlined in the job aid seen in Figure 8 and used it to define the potential hurt level worst-case (PHLwc) for the sample set. For the sample set analysis, the PHLwc was the scenario that is most probable and produced the highest hurt level without additional hazards that were not present at the time of the incident. For example, if a scenario includes the use of oxidizers or flammables, but there was no sparking or flame hazards typically involved in the work or area nor anticipated to be present, a practitioner would not define a scenario where ignition occurs causing a significant hurt level as probable.

Table 5. Hurl Level Definitions. Adapted from [15].

Hurt Severity Level	Duration	Definition	Examples
Fatality (4)	N/A	Fatality	- Fatality or Multiple Fatalities
Severe Hurt (3)	Years to Lifetime	Injury or illness causing severe physical body damage; probable long term or significant life-altering complications	- Amputation - Significant third-degree burns - Loss / Impairment of organ functions - Severe to complete loss of hearing - Severe or total blindness
Moderate Hurt (2)	Weeks to Months	Injury or illness causing significant physical body damage; reasonable to heal without life-altering complications in a moderate period	- Fractures, loss of tooth/teeth - Significant lacerations - Partial / single digit amputations - Significant second-degree burns - Moderate hearing loss
Minor Hurt (1)	Minutes to Days	Injury or illness causing minor physical body damage; reasonable to heal without significant life-altering complications in a brief period	- Minor lacerations that bleed freely - Minor chipping of tooth/teeth - Skin rash / burn from chemical / non-aqueous fluids - Confirmed slight to mild hearing loss - Mild corneal abrasion
No Hurt (0)	N/A	N/A	N/A

Because multiple scenarios are often considered to derive the worst-case scenario the practitioner has to qualitatively screen the scenarios to find the one where the value of the product of the hurt level, based on the definitions in Table 5, and probability of the event is the highest. Figure 8 shows a visualization of the mental model a practitioner goes through when determining the worst-case scenario for an incident. The product of the hurt level and likelihood of the scenario for each scenario generate various possible outcomes. Therefore, the hurt level of this scenario is the scenario with the highest potential hurt level for the worst-case scenario (PHLwc) and is used later in the pSIF model.

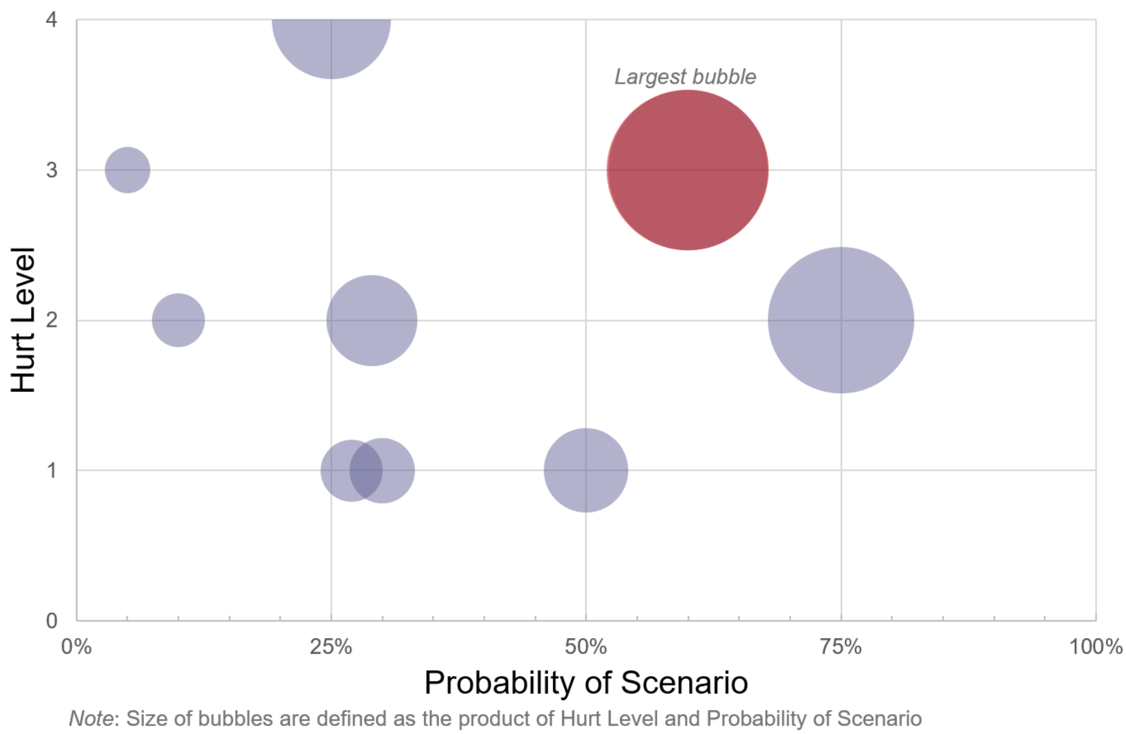

Figure 8. Hurt Level versus Scenario Probability: A Bubble Chart Analysis

E. PRIMARY VARIABLE

In conceptualizing the worst-case scenario, the practitioner identifies the primary variable, that if changed, would result in the worst case outcome. Typically, a practitioner assigned a qualitative value to the primary variable based on prior experience, any FMEA [36], or a documented method and classify it under one of the following values from most probable to least probable: Always, Often, Sometimes, Rarely, Not at All. The most feasible method to implement was determined to be a qualitative determination as performing a quantitative probability analysis on the incidents was not always possible for incidents. The qualitative nature of assessing probability is also a clear limitation of this pSIF model.

The model requires the practitioner to define the primary variable (V) that would have led to the PHLwc. For example, in a scenario where a machinist is injured while using a lathe, a primary variable that would have increased the hurt level could have been body placement. Note that practitioners were advised to not increase hazards or introduce new hazards. Building off the lathe example, if the work typically includes working with steel, the practitioner will not replace the material with depleted uranium, or other heavy metals if it is not typical to the work in the shop. Therefore, V provides the variable in which a worst-case scenario would occur but might not have due to random-chance.

It is also important to note that decreasing the primary variable would not reduce the potential hurt level altogether. Which is why determining control effectiveness was key for the pSIF model.

F. DETERMINING CONTROL EFFECTIVENESS

Typically, there are multiple layers of control in an incident as defined by the Swiss Cheese model [32]. Therefore, removal of a variable does not remove hurt level altogether. For example, latex gloves may prevent or mitigate hazardous chemical exposure but may do little in preventing laceration if working with sharps. So, practitioners must consider hazards that are present in the work environment of the incident, such as de-energized equipment that may be otherwise energized due to a failure in administrative controls, or a pressure system failure due to use of an incorrect rupture disc. Such scenarios may come up in discussion of the work control process and the level of rigor may depend on the hazards involved [35] or looking at the failure mode and effects analysis but looking at the details of a scenario post-incident relied on methods similar to the Haddon matrix [37]. Either way, it is a systems approach to safety that must take into account various hazards and the controls in place that mitigate those risks.

Hazard controls in place contribute to the probability of a worst-case event. Even in a criticality event, fissile material controls will influence a criticality event outcome [32] and the severity of an incident. A simplified diagram illustrates the relationship between incident severity and the primary variable as well as incident probability and hazard controls, is shown in Figure 9. The relationship in hazard controls and incident probability is not novel; James Reason developed what is now referred to as the Swiss Cheese model in the late 1990s, theorizing that adverse events usually result from many controls failing to prevent the incident, and that each layer of control acts as separate barriers against the incident similar to attempting to pass through slices of Swiss cheese

[33]. The Swiss Cheese model is also used in risk management, and flight mishap analysis [6] as it is often the failure of multiple sub-systems within a system that leads to the outcome. Additionally, nuclear operations will employ defense in depth in concepts of security and in human error [34]. But the model must be simple enough to be deployed institutionally.

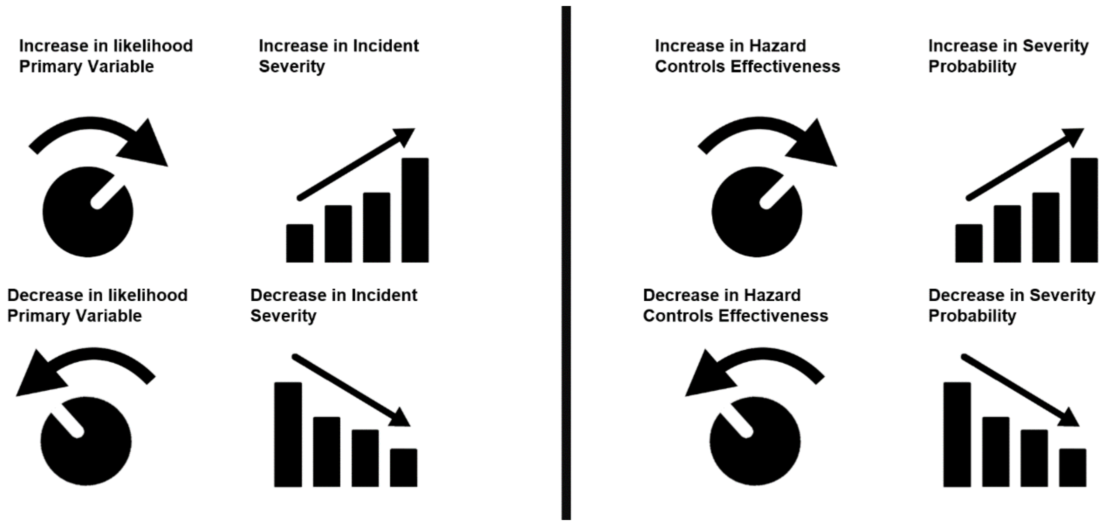

Figure 9. Incident Severity and Event Probability Correlation

For over the last fifty years the National Institute for Occupational Safety and Health (NIOSH) [38] and the Occupational Safety and Health Administration has advocated the hierarchy of controls approach to emphasize that engineered and more permanent controls are preferred as they are more effective in mitigating risk, a copy of this illustration is seen in Figure 10 [39]. The controls in each incident were evaluated based on overall effectiveness, using the hierarchy as a guide, in their effectiveness in mitigating the worst-case outcome.

If the data element included details on controls such as personal protective equipment (PPE), then it was understood that PPE was a control at the time of incident, regardless of use. However, if PPE is the only control, then it was further scrutinized as it is generally seen as the least effective control in the hierarchy of controls. Similarly to the determination of the primary variable (V) a qualitative approach is used to categorize the effectiveness of the controls ranging from least effective to most effective: No Controls, Some Ineffective/Inconsistent Controls, Controls Reduce Severity, Controls Likely Prevents PHLwc, and Controls Always Effective. The score is known as the control effectiveness score (C).

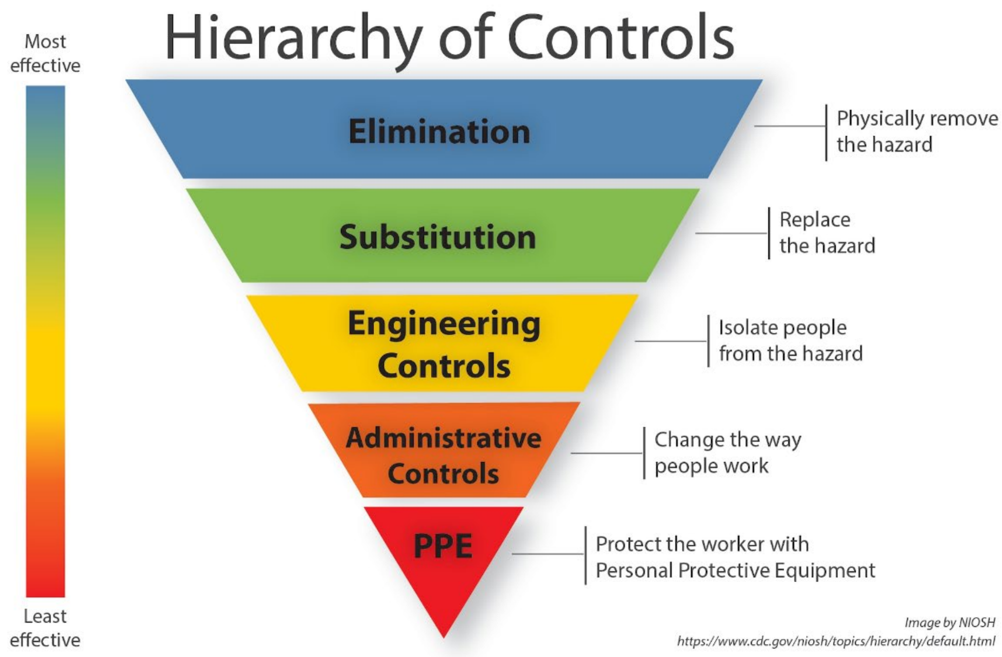

Figure 10. Hierarchy of Controls. Source: [38].

G. DETERMINING FINAL PSIF SCORE

Practitioners at the final stage of framework have determined the PHLwc, V, and C. Armed with all the variables and the job-aid in Figure 7, there is now sufficient data to determine the final score for the incident. Determining the V/C score is done through the use of the matrix in Table 6. Output of the matrix, V/C score, is then used with the PHLwc to produce the final potential hurt level (PHLf) score.

Table 6. V/C Matrix. Source: [15].

	Likelihood of Primary Variable (V)				
Effectiveness of Hazard Controls (C)	**Always**	**Often**	**Sometimes**	**Rarely**	**Not at all**
No Controls	5	4	3	2	1
Some Ineffective/ Inconsistent Controls	4	4	3	2	1
Controls reduce severity	3	3	2	2	1
Controls Likely Prevents PHLwc	2	2	2	1	1
Controls Always Effective	1	1	1	1	1

Using Table 7, the PHLf is determined by looking at the PHLwc and V/C score. Those with a PHLwc score of Severe (3) and a V/C score of at least three are treated as pSIF. But also used a level of rigor approach such that those with a V/C score of 2 and PHLwc of a Fatality (4) still generated a pSIF classification. Those with a PHLf score of 4 are treated as those with an even higher potential for a SIF incident. This is either because the likelihood of it occurring is high, or because the controls in place to prevent such an incident are either not effective or not in place. These incidents should be given a higher priority in mitigating the pSIF scenario. Inherently, those incidents that are given a higher priority would undergo a more rigorous investigation leading to the corrective actions or institutional response. The institutional response was also categorized as part of the case study and is discussed in the next section.

Table 7. PHLf Score Matrix. Source: [15].

	PHLwc				
V/C Score	**Fatality (4)**	**Severe (3)**	**Moderate (2)**	**Minor (1)**	**No Hurt**
(5)	4	3	2	1	0
(4)	4	3	2	1	0
(3)	4	3	2	1	0
(2)	3	2	1	0	0
(1)	2	1	0	0	0

H. CATEGORIZING INSTITUTIONAL RESPONSE

As part of the case study, the institutional response to the actual incidents were evaluated to determine if there was any existing correlation between incidents with a higher potential for serious injury or fatality and the institutional response. Organizations will typically benefit from the prioritization of mitigation efforts to address risk that will cause the most severe harm to the organization's workforce [40]. At Lawrence Livermore National Laboratory there are several ways to generate corrective actions to mitigate institutional risk.

The responses to the organization were tied to the effect the action had at the institutional level and how many participants were associated. It is a measure of resources used in an incident response but is in no way definitive. The scores were broken down into five separate categories, each increasing in rigor, and was called the preventative action score (PAS) for simplicity. Table 8 explains the level of rigor of each PAS category. Each of the incidents from the representative sample were given a score based on the criteria. However, because PAS only looks at a specific type of institutional response, it may not holistically review the resources allocated in a corrective action and is a limitation of this type of categorization.

Table 8. Preventative Action Scores and Definition

Preventative Action Score	Institutional response	Example of Response	Number of People Affected
0	No Response	No Corrective Actions	No Affected Parties
1	Individual Involved in Incident	Verbal counseling	One person affected
2	Group or Team Level	Lessons learned shared at group meeting	Two to ten people affected
3	Organizational or Facility Change	Self-audit across one or multiple facilities	Ten to hundreds of people affected
4	Institutional Change / Hazard Elimination	Institutional policy on work with specific materials	Hundreds of people to total workforce population

III. RESULTS OF CASE STUDY

A. FINDINGS

The goal of the case study was to assess the feasibility of the pSIF framework against the representative sample outlined in Chapter II. The SI score was only used to indicate the feasibility of adapting the proposed framework to existing institutional procedures. There were 48 incidents that had an SI score of less than or equal to 2, and the random sample was meant to represent the variance in the dataset with a 95% confidence level with a 5% margin of error. It was concluded that the current process captures enough information to implement the framework. The only suggestion would be to purposefully track causal information or additional information about the hazards as those two categories were the least likely to be included in an incident report. This would increase understanding of an incident to better understand the highest potential severity of an incident outcome.

However, the relatively high percentage of samples that had information about the controls in place and environmental factors suggest that the process is sufficient to create incident information to implement the model during postmortems. Framework adoption is not expected to greatly change the current process or methodology. However, from a human performance perspective, framework adoption may encourage case managers and investigators to note the cause and hazards of an incident, potentially improving data quality. However, tracking of these data elements is necessary to see any long-term impacts.

Several practitioners evaluated the incidents through the decision tree described in Figure 7. The final distribution of the final potential serious injury or fatality (PHLf) score, illustrated in Figure 10, indicates that only 8 incidents; less than 1% of the total – had a high potential for a SIF. An in-depth statistical analysis of the distribution is done in Chapter IV. and Appendix B. It is also worth noting that 33% of the cases had a very-low potential, PHLf score of 0, while 52% of the cases had a low potential, PHLf score of 1.

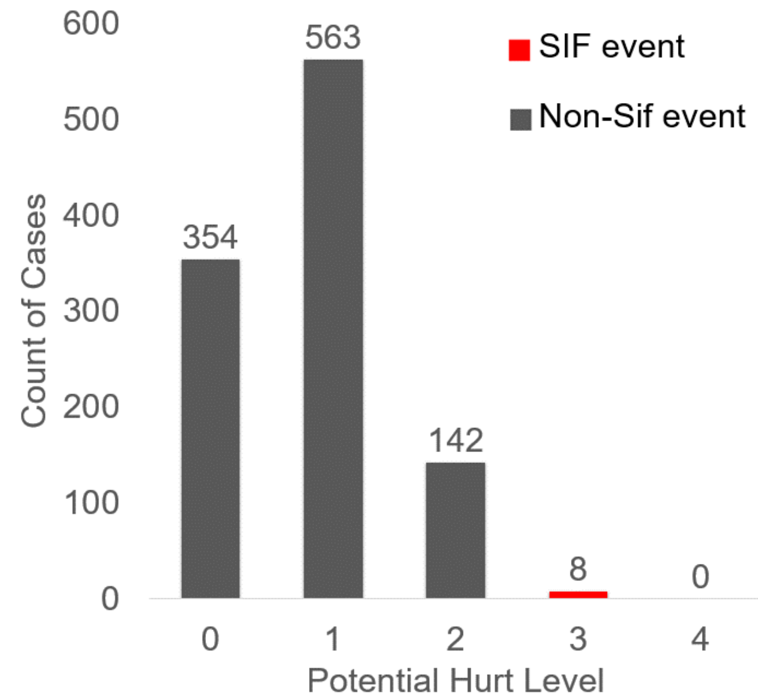

Figure 11. Distribution of PHLf Scores from the Case Study. Source: [15].

The distribution of PHLF scores does not follow the distributions seen in Heinrich's safety triangle [8]. This could also be due to the fact that the case study only included recorded incidents. Limitations of the dataset include the exclusion of unsafe acts or behaviors, incidents that did not result in an injury, and those not formally reported in the institution's near miss program. The distribution does mimic distributions seen from some private industry models [14] which suggests that only a small percentage of unsafe behaviors have the potential to cause serious injury. Table 9 presents the distribution of incidents grouped by accident type and PHLf score. This data can be used to inform relevant computational modeling.

Table 9. Breakdown of Accident Type by PHLf Score

Accident Type (Group)	PHLf Score	
	2	3
Assault	1	
Bodily Reaction	10	
Caught in Equipment, Materials, or Machinery	3	1
Contact	7	1
Fall (all types)	7	
Ingestion of Substance	2	
Inhalation of Substance	6	
Noise Exposure	34	1
Not Specific	1	2
Overexertion	12	2
Radiation	1	
Reaction when surprised, frightened, startled	1	
Repetitive Motion	43	
Rub or Abrasion	1	
Self-Inflicted Injury	1	
Stationary Injury	2	
Struck by Object	7	
Vehicular Accident	3	1
Grand Total	142	8

Ideally, a review of the incidents with a high potential to cause life altering ailments is of high interest to the organization. Additionally, any significant trend or uptick in incidents of a specific accident type should be regularly reviewed. Although incidents with a PHLf score of 3 were statistically insignificant, examining those with a score of 2 is also valuable. The top three categories were then found to be noise exposure, overexertion, and repetitive motion and at 89 incidents made up 62% of the incidents.

Because overexertion was already one of the leading accident types from the data set, this suggests that randomly sampling the whole data set could have inadvertently biased the data towards those incidents with higher incident numbers. The other top categories were evaluated to understand the long-term implication of the occupational hazard.

Understanding the long-term implications of these incidents helps allocate resources effectively within organizations. Studies showed that about 22 million workers are exposed to hazardous noise at work on a yearly basis and that hearing loss and tinnitus affects 1 in 8 people in the working population where occupational exposures caused 25% of those incidents [41]. It is also important to note that noise exposure affects industries from the military [42] to private industry [40].

The next categories – overexertion and repetitive motion – can lead to lifelong musculoskeletal disorders such as chronic back pain, arthritis, and can result in disability and work loss [43]. Since eighteen of the accident types make up less than 40% of the rest of accidents, it is unlikely that it shows a statistically reliable way of predicting future incidents in a computational model. Lastly, because the focus of the framework is potential high severity events, the count of accident type may not be the best measure for resource allocation.

Another important aspect of the case study was the Preventative Action Score (PAS), which was used to determine an aspect of LLNL's response to incidents. Institutional response can influence future outcomes, similarly to how risk is mitigated in an integrated safety management system by relying on feedback and continuous improvement [22], [23], [35]. Research has shown that simulations can identify corrective management actions that can reduce risk in the construction industry [44]. Unfortunately, the diverse range of hazards apparent in research and development complicate implementing sociotechnical simulations. Additionally, the PAS definitions were not tied to monetary value, which makes it a difficult measure of financial resources allocated in incident response. A future approach would take into consideration the numerous factors that is influenced by institutional response, such as resources used, engineering methods implemented, and labor hours in addition to other actions.

Figure 12 illustrates the PAS distribution by PHLf score. Since both PAS and PHLf are categorical, performing a linear regression would be inappropriate as it would require violating assumptions regarding the continuity and homoscedasticity of the data which may cause unintentional bias or skewing [45]. Instead ordered logistics regression can be used to see if there is any correlation between the PHLf and PAS [46]. The analysis used the OrderedModel class from the Python module, statsmodels [47], and pandas [48]. Additional models were also explored and are explained in greater detail in later chapters.

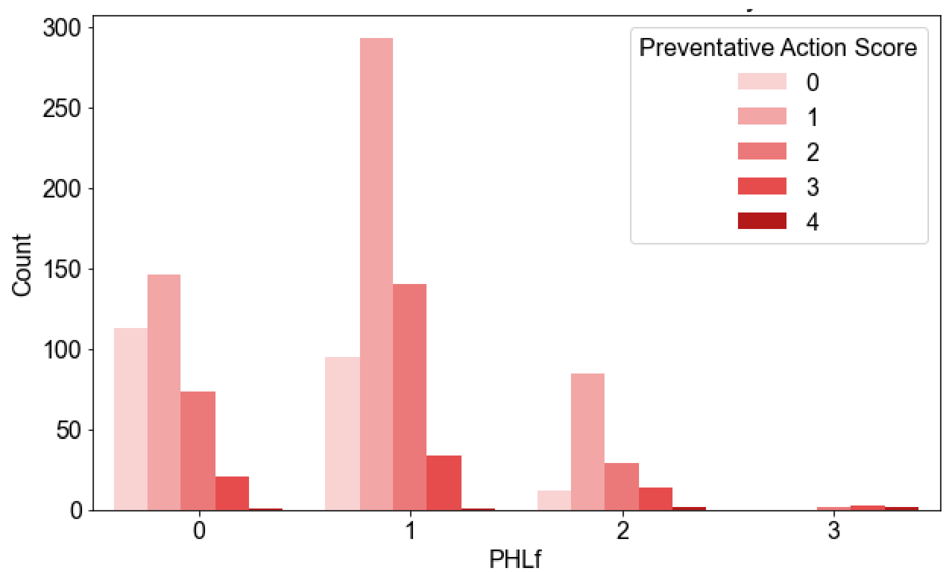

Figure 12. Distribution of Average Preventative Action Score by PHLf

Overall, the case study demonstrated that the current injury and illness incident reporting process is adequate to support a potential severity framework. While some incidents had a higher potential for sever outcomes, the majority of incidents were not classified as pSIF incidents. Statistical evidence indicated that LLNL generally responded more robustly to incidents with higher potential [15]. This is important to note as the goal of implementing the framework would be to increase response of higher severity incidents to better allocate resources. However, future research is needed to better understand incidents with a PHLf of three or higher as the model may suffer from under dispersion due to the limited number of incidents with a PHLf score of 3 or 4.

B. LIMITATIONS OF THE CASE STUDY

If the case study is to be used as the basis for any computational modeling, it is necessary to transparently acknowledge its limitations. Especially as these factors may impact any operational implementation. Some pSIF models from private industry [13], [14], [16] and Heinrich's safety triangle [15] described in Section I. B. pSIF Models in Private Industry considered near-miss and unsafe behaviors. However, the data set used for the case study did not include near-miss information, such as that available at the Department of Energy Office of Environment, Health, Safety & Security lessons learned website DOE OPEXShare [50]. This information could have provided near-miss opportunities that had a higher potential severity, helping calibrate the framework and dealing with issues of underrepresented data with higher PHLf scores.

Additionally, the EEI pSIF model recognized the limitations in existing methods of classifying safety incidents. This includes subjectivity in the assessment, generalized conditions that may not indicate SIF potential, the broad use of an "other" category by analyst, and no explicit consideration of physical controls [16]. The case study attempted to mitigate the identified limitations by minimizing the use of an "other" category as part of the injury and illness reporting process. Furthermore, the framework minimized the ability to extrapolate generalized conditions. This was done to mitigate inadvertently misclassifying an incident. For example, personnel working in a glovebox environment with radioactive material does not imply that a fall in the same lab space poses a radiation concern.

The practitioners also considered the presence and absence of controls. The case study, due to the broad nature of research and development at the national laboratory, could not eliminate subjectivity in the probabilistic determination of the primary variable or controls effectiveness while assessing the incidents and remains a known limitation of the framework. Providing quantitative measures for specific hazards would mitigate this risk but requires comprehensive methods of quantifying incident outcomes across a spectrum of work.

Finally, the limited number of incidents with a high potential for a serious injury or fatality resulted in a need to implement a computational model to assist in understanding framework implementation. Despite these limitations, the case study formed foundational insights that could have implications in the implementation to other management systems.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. A COMPUTATIONAL MODEL OF THE FRAMEWORK

A simple computational model was developed to simulate the potential serious injury or fatality framework described in Section II. Subsection C. The objective was to perform correlation analysis between the variables used in the pSIF model to ensure it functioned as expected. Specifically, to see if the model informs incidents with high hurt potential and high probability, combined with low hazard controls are effectively classified as pSIF incidents. Simulated data from this computational model serves as a control for the statistical modeling of institutional response to the events. Modeling provides controlled parameters and the ability to simulate datasets where data may be difficult or rare. Figure 13 illustrates how the pSIF statistical model will inform institutional response and ultimately how these models will inform operational use of the model.

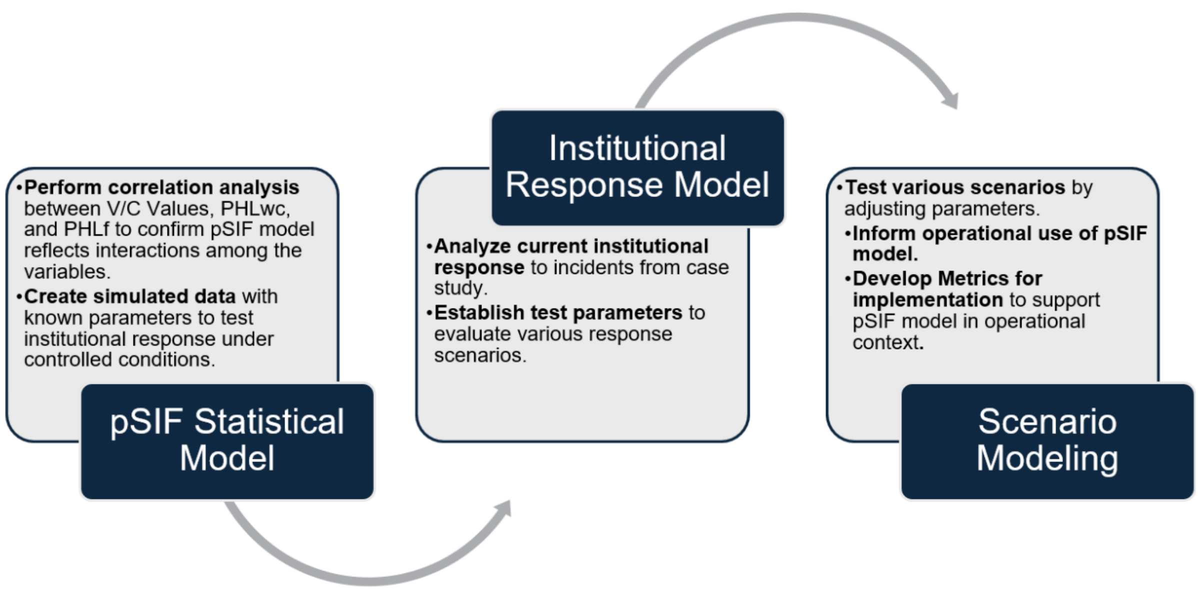

Figure 13. Workflow for pSIF Statistical Model and Institutional Response Analysis

The computational model seeks to analyze institutional response, as done in the case study, by performing a correlation analysis on the case study data. This model simulated institutional response. Various scenarios were evaluated by adjusting the parameters.

A scenario where…

- The potential severity of an incident is commensurate with the institutional response effort, the assumed ideal scenario.

- The potential severity of an incident is based on the distribution found in the case study, without framework implementation.

- The institutional response is completely random with no consideration for incident severity.

In performing the various scenarios, the computational model was used to evaluate potential improvement in resource allocation. The model also informed the development of meaningful metrics necessary for tracking any long-term improvement through implementation. The computational model, plots, and testing was built in Python using several libraries such as Matplotlib [57], Numpy [58], Pandas [48], Seaborn [59], Scikit-Learn [60], and Statsmodels [47]. The models and simulations were made in Python to leverage open-source software and were saved in a Jupyter notebook [61] for repeatability and sharing.

Several statistical models were tested with mixed results. The case study looked at a generalized linear model (GLM) with a Poisson distribution and an ordinal logistic regression model. Each model type has its own set of assumptions and limitations. The results and detailed development of the GLM with a Poisson distribution and ordinal logistic regression model tested is explained in Appendix B. pSIF Statistical Models.

A. SIMULATING THE CASE STUDY INCIDENT DISTRIBUTION

Injury and illness (I&I) data is typically very sparse and often require datasets from other industries and organizations for meaningful analysis [56]. I&I data is also typically difficult to obtain due to concerns with Personal Identifiable Information (PII). For the case study, the incidents spanned over a decade at LLNL in order to mitigate the difficulty in gathering a large enough dataset. It was also important for the case study to focus on hazards relevant to LLNL, which differ from those in industries like oil and gas.

Simulating data is useful for several reasons. It allows creation of incidents without having to rely on actual occurrence. This is especially important for rare incidents like SIFs. Additionally, this work aimed to evaluate resource allocation differences in implementing the model. The simulated data also offers ways to emulate institutional responses for a wide distribution of incidents and help aide operational implementation.

First, the data from the case study was loaded from an Excel file and into a Pandas DataFrame, including the values for PHLwc, V, C, V/C matrix, PAS, and PHLf for each incident. The data contained categorical features that needed encoding for statistical modeling. Because the data was ordinal with inherent order (e.g., increasing control effectiveness), the values were assigned based on rank. Figure 14 shows the value, category, and encoding for each variable. Table 11 presents the descriptive statistics of the encoded data. The data indicated that the PHLf, PHLwc, C, and the PAS are skewed towards the lower half of the ranking. Notably, the PHLf score only has a maximum score of three, where the maximum rank for the variable is four. Simulating the data allows the ability to create datasets where there is a PHLf score of four and what might have been the possible PHLwc, V, and C values to produce the score.

Increasing Rank →

Variable	**Potential Hurt Level Worst-Case (PHLwc)**				
Category	No Hurt	Minor Hurt	Moderate Hurt	Severe Hurt	Fatality
Encoded Value	0	1	2	3	4

Variable	**Primary Variable (V)**				
Category	Not at All	Rarely	Sometimes	Often	Always
Encoded Value	0	1	2	3	4

Variable	**Controls Effectiveness (C)**				
Category	No Hazards Controls	Some ineffective/ inconsistent controls	Controls reduce severity, but not prevent PHLwc	Sufficient controls likely prevents PHLwc	Thorough controls always effective
Encoded Value	0	1	2	3	4

Figure 14. Encoded Ordinal Variables in Increasing Rank

Several models were tested to determine any statistically significant correlation between the PHLf score and PHLwc, V, C, and V/C. Because higher PHLwc and V/C scores created a higher PHLf outcome from the framework, the modeling aims to fit the distribution of the data from the case study.

Table 10. Descriptive Statistics of Encoded Case Study Data

	PHLwc	**V**	**C**	**V/C**	**PHLf**
Mean	1.71	2.66	1.16	2.11	0.82
Standard Deviation	0.64	0.58	0.99	0.43	0.68
Minimum Value	0	0	0	0	0
Maximum Value	4	4	4	4	3

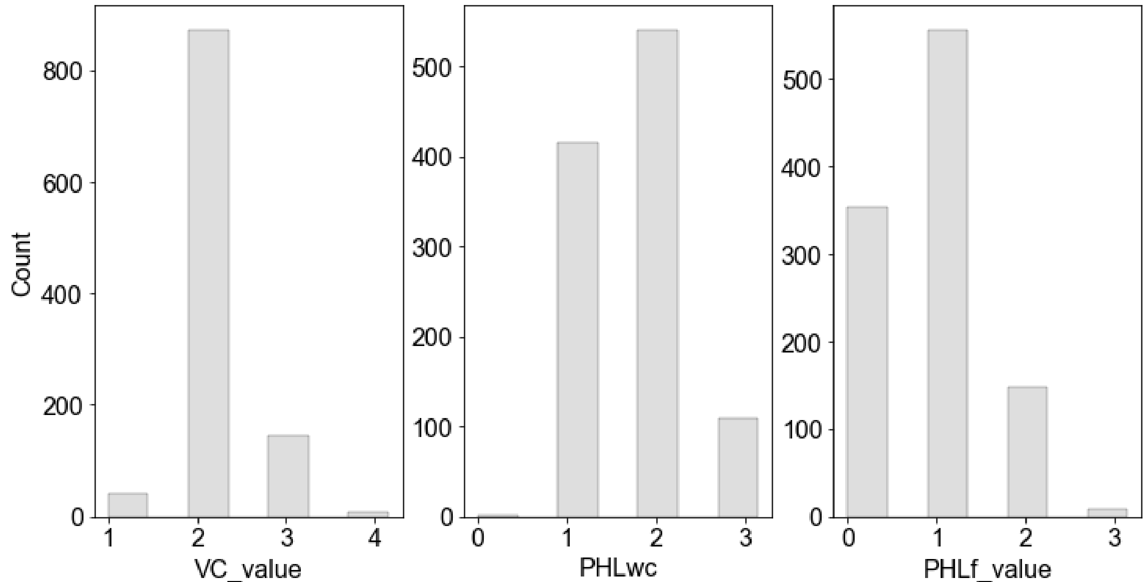

Figure 15. Distribution of Case Study Values

B. OVERVIEW OF STATISTICAL METHODS EVALUATED

Understanding the key assumptions and limitations of each method is crucial for building robust and flexible computational models to simulate the necessary data. For this reason, two statistical models were used for the work based on their use in evaluating categorical data [45].

When selecting the appropriate statistical modeling technique, various assumptions and limitations were considered. These factors influence model performance and validity in operational use. Table 12 provides a summary of key assumptions, limitations, and benefits of each model evaluated.

Table 11. Statistical Models Overview. Adapted from [45], [46].

Model	Key Assumptions	Limitations	Benefits
Ordinal Logistic Regression	Ordinal dependent variables (i.e., the ratings have an order as discussed in the framework).	The relationship between each pair of outcome categories is consistent across thresholds.	Specifically designed nuanced analysis of ordered categories.

GLM with Poisson Distribution	The data is discrete (e.g., number of incidents). Assumes that the data follows a Poisson distribution.	Assumes that the mean and variance of the data is always equal.	Effective for modeling count data.

The goodness of fit for each statistical model was evaluated to understanding the relationship of the PHLf score and its distribution. Appendix B. pSIF Statistical Models explains in detail how the ordinal logistic regression model and GLM with Poisson distribution model were created and how they work.

C. COMPARING RESULTS OF THE PSIF MODELING

The Statsmodels [47] package from Python was used to create the pSIF statistical model. There are several methods in evaluating how well the model's goodness of fit. Three variables were used to compare the goodness of fit, including the Log-Likelihood Function (LLF), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) [45]. Details of how the LLF, AIC, and BIC were calculated are described in detail in Appendix B. Section C. Variables to Compare Goodness of Fit.

The LLF measures how well a statistical model explains observed data. It is the log of the likelihood of the observed data given the parameters and a probability distribution. The LLF is also calculated differently for each model because of how the cumulative probabilities is calculated. For an ordinal logistic regression model, it involves calculating cumulative probabilities for each PHLf score, with the LLF being the sum of the log of these probabilities weighted by an indicator variable. For a GLM with a Poisson distribution, the LLF is derived from the Poisson probability distribution.

Using the formulas detailed in Appendix B. Section C. Variables to Compare Goodness of Fit, the LLF, AIC, and BIC were programmatically calculated in a Python script. The results of the model are seen in Table 12. In comparing model goodness of fit, the goal is to minimize each of the different criteria and there is a large discrepancy between the values produced by the ordinal logistic regression model and the generalized linear model with Poisson distribution. Therefore, the ordinal logistic regression model seems to be better at modeling the data from the case study. However, to ensure that this is the case, both models were tested to create a simulated data set. Ideally, the model must be robust enough to change parameters to deal with changes in future incident distributions.

Table 12. Descriptive Statistics of pSIF Statistical Models

Criteria	Ordered Model	GLM with Poisson Distribution
LLF	-1.4e-4	-881
AIC	14	1771
BIC	49	1796

Comparatively, the ordered model quantitatively had a better goodness of fit. With this in mind, the result of the model was analyzed and both statistical models showed a statistically significant correlation between the PHLf value and the PHLwc and V/C value as seen in Table 14. However, a significant limitation of the ordered model was in its ability to effectively predict values that would generate a PHLf score of 4, as there were none that occurred during the case study. This severe limitation raises concerns on the ability to generalize to new data and was determined to be ineffective in simulating the incidents needed for testing. But both results confirm that the pSIF model is behaving as expected, with a clear correlation between the PHLf value and dependent variables. There is no reason to believe that the V or C score alone have any correlation with the PHLf score as both scores contribute to the V/C score.

Table 13. Summary of pSIF Model Coefficients

	Model	
Coefficients	Ordered Model	GLM with Poisson Distribution
PHLwc	393.7213	1.1627
V	-157.4856	0.0995
C	-151.3317	0.0018
V/C	60.6603	1.0341

D. SIMULATING PSIF DATA

Because the case study is based on a random sample of over ten years, it may be a good indicator of what the distribution of injury and illness incidents may occur over time. For the thesis, the goal was to also synthesize pSIF data to model institutional response under various scenarios when implementing the framework operationally. Creating metrics is important in tracking framework influence in the organization and can indicate the efficacy of the framework [56]. The next section details simulating the pSIF data used to model institutional response. The simulated data allows to model response to incidents with a PHLf score of 4, which there were none in the case study data. Because the ordered logistic regression model was fitted to the only available categorical values available, the GLM with Poisson distribution was used as the basis to simulate data that will be used for testing institutional response against a dataset where the parameters are well characterized and known.

The model provided the statistical parameters to provide the distribution of PHLf scores. Because the PHLf score was mostly influenced by the PHLwc and the V/C score, those two variables will be used to generate the synthetic data. Figure16 is a screenshot of the code that was used to generate synthetic pSIF data that will be used later on. The code works by generating PHLwc and V/C scores based on the descriptive statistics of those variables from the case study data. It then clips the data to ensure it meets the constraints of the framework and generates the appropriate number of variables as defined by the sample size desired. Next, the function calculates the λ-value which defines the shape parameter of the Poisson-distributed outcomes. The shape parameter takes the place of using thresholds in an ordinal logistic regression model and uses this to generate values based on discrete probabilities. It then uses the λ-value to generate the PHLf scores and is converted to discrete values that meet the constraint of the framework.

In [112]:

```
# Function that generates simulated sample sizes based on the distribution of the model or given beta values
def simulate_data(sample_size, beta_PHLwc, beta_VC_value, mean_PHLwc, std_PHLwc, mean_VC_value, std_VC_value):
    # Using the mean and standard deviation from the model
    PHLwc_values = np.random.normal(loc=mean_PHLwc, scale=std_PHLwc, size=sample_size).round(0)
    PHLwc_values = np.clip(PHLwc_values,0,4)
    VC_value_values = np.random.normal(loc=mean_VC_value, scale=std_VC_value, size=sample_size).round(0)
    VC_value_values = np.clip(VC_value_values,0,4)
    # Calculate the lambda for the predictor values
    lambda_values = np.exp(intercept + beta_PHLwc * PHLwc_values + beta_VC_value * VC_value_values)

    # Generate Poisson-distributed outcomes
    simulated_PHLf_values = np.random.poisson(lam=lambda_values)

    # Apply constraints to ensure values do not exceed framework values
    simulated_PHLf_values = np.round(simulated_PHLf_values)
    simulated_PHLf_values = np.clip(simulated_PHLf_values, 0, 4)

    return pd.DataFrame({
        'PHLwc': PHLwc_values,
        'VC_value': VC_value_values,
        'PHLf_value': simulated_PHLf_values
    })
```

Figure 16.   Code Used to Generate Simulated PHLf Values

Three separate sample sizes were created simulating 1,000, 1,100, and 1,200 incidents respectively. Figure 17 illustrates the results for three sample data sets that were created to compare to the case study data and evaluate the fit of the model.

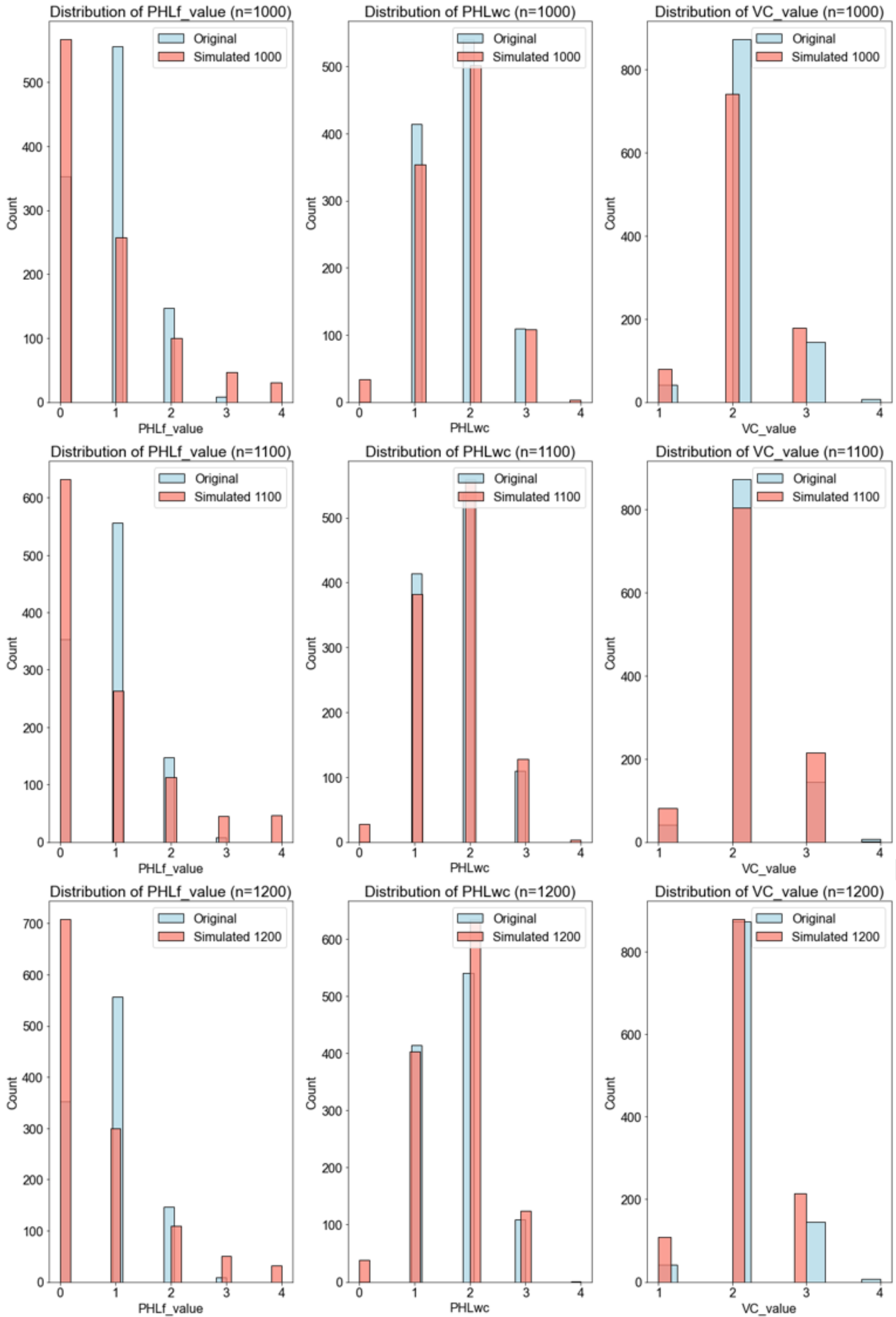

Figure 17. Results of GLM with Poisson Distribution Simulation

E. SIMULATING ORGANIZATION RESPONSE

Now that the data can be modeled to encompass the full range of PHLf scores, a separate statistical model was created to simulate current institutional responses to the incidents. This model was developed to understand the relationship between the PHLf scores of the case study and the institutional response to the incident defined as the Preventative Action Score (PAS). However, the model will need to be flexible in order to change the parameters to represent the various scenarios outlined in the beginning of this chapter to simulate different institutional responses.

In an ideal scenario, the institution would increase response to incidents with a higher PHLf score. However, this may not always be ideal due to the definitions of the PAS categories as outlined in Chapter II. Section G.

There are scenarios where a lower institutional response can mitigate incidents with higher PHLf scores. For example, if an incident occurs relating to explosives occurs in a facility with a high PHLf score, but does not affect the operations of other facilities, the institutional response may be more localized and generate a PAS value that is localized to the facility. In this way, adequate response was considered using an appropriate graded approach.

Additionally, injury and illness incidents are highly variable, which is why a comprehensive safety management system, and risk assessments are incredibly important. The importance of acceptable risk and its industrial safety guidelines have been outlined by various organizations, including the International Atomic Energy Agency (IAEA) [70], [71] which outlines specifics industrial safety guidelines and risk aggregation for nuclear facilities. The IAEA has also outlined a concept similar to the pSIF model but calls those with a high-potential for harm (HiPo) incidents, also generating near miss analysis for such instances [70].

1.Modeling the Relationship Between PHLf Score and PAS

The ordinal regression was thought to be appropriate because the preventative action score is categorical with a natural order, but not with a known interval between categories. In this case, the PHLf is the predictor variable to estimate the probability of each preventative action score by maximizing the probability of observing the given data as a function of model parameters in a likelihood function, $L(\theta)$[46]. The maximum likelihood (ML) is an estimate that the parameter value maximizes this function and using numerical optimization techniques, the fit method in the statsmodels library iterates through various model parameters until the algorithm has reached convergence. For an ordered logistic regression model, the likelihood function can be expressed as $L(\theta) = \prod_{i=1}^{n} P(y_i \mid x_i, \theta)$ where $P(y_i \mid x, \theta)$is the probability of observing a preventative action score, y_i, given a PHLf, x_i, and where n is the number of observations [46].

The log-likelihood function was used to ensure the predicted probabilities are always between 0 and 1 and for any non-linear relationships that may have been found between the PHLf and preventative action score and is expressed as $\ln L(\theta) = \sum_{i=1}^{n} \ln P(y_i \mid x_i, \theta)$. The optimization method used was the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [49]. The code and results are in Figure 18, where severity is the PHLf, and response was the preventative action score.

```python
# Create a DataFrame
df = pd.DataFrame(data)
''' Define the independent variable (predictor/PHLf)
and the dependent variable (Preventative Action Score)'''
X = df[['severity']]
y = df['response']
# Fit the ordinal logistic regression model
model = OrderedModel(y, X, distr='logit') #Logit function
result = model.fit(method='bfgs') #Use Broyden-Fletcher-Goldfarb-Shanno
print(result.summary()) #print summary
```

```
Optimization terminated successfully.
         Current function value: 1.208498
         Iterations: 17
         Function evaluations: 18
         Gradient evaluations: 18
                             OrderedModel Results
==============================================================================
Dep. Variable:               response   Log-Likelihood:                -1288.3
Model:                   OrderedModel   AIC:                             2587.
Method:            Maximum Likelihood   BIC:                             2611.
Date:                Wed, 24 Jul 2024
Time:                        21:24:14
No. Observations:                1066
Df Residuals:                    1061
Df Model:                           1
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
severity       0.4911      0.087      5.621      0.000       0.320       0.662
0/1           -0.9644      0.100     -9.602      0.000      -1.161      -0.768
1/2            0.8028      0.038     21.139      0.000       0.728       0.877
2/3            0.5389      0.063      8.537      0.000       0.415       0.663
3/4            0.9737      0.149      6.527      0.000       0.681       1.266
==============================================================================
```

Figure 18. Logistic Regression Results

The coefficient for the PHLf was estimated at 0.4911 with a standard error of 0.087, yielding a z-value of 5.31 and a p-value that is statistically significant at 0.001 (<0.05). The results suggest that the increase in PHLf score did have a statistically significant correlation with higher preventative action scores. Future work is still needed to understand those cases with a PHLf of three or higher and it is possible that the model suffers from under dispersion due to the low amount of incidents reviewed with a PHLf of 3 or 4.

The methods used to model the relationship between the PHLf score, and the PAS are outlined in Appendix B. pSIF Statistical Models. However, in this case, the PHLf score was used as the predictor value and the PAS was defined as the target variable. Table 15 shows some descriptive statistics of the case study data for the PHLf score and

PAS. Figure 11 also shows the distribution of the PAS for each PHLf category, which is useful in understanding how the institution has responded to incidents of varying severity.

Table 14. Descriptive Statistics of Case Study Data for PHLf and PAS

	PHLf	**PAS**
Mean	0.82	1.18
Standard Deviation	0.68	0.85
Minimum Value	0	0
Maximum Value	3	4

Additionally, Figure 19 shows the code that was used to create GLM with Poisson distribution model along with summary of the results. The model is simpler in that only one variable, the PHLf sore, is used as the predictor for the target, PAS. This was done to more easily change the coefficient to model distributions outside of what was found in the case study.

```
In [173]:
X = case_study_data[['PHLf_value']]
y = case_study_data['PAS']

# Add a constant to the model (intercept)
X = sm.add_constant(X)

# Fit the Poisson regression model
model = sm.GLM(y, X, family=sm.families.Poisson()).fit()

# Print the summary
print('GLM Results w/ PHLf_value as the sole predictor', end='\n\n')
print(model.summary(), end='\n\n')
```

```
GLM Results w/ PHLf_value as the sole predictor

                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                    PAS   No. Observations:                 1064
Model:                            GLM   Df Residuals:                     1062
Model Family:                 Poisson   Df Model:                            1
Link Function:                    Log   Scale:                          1.0000
Method:                          IRLS   Log-Likelihood:                -1355.5
Date:                Wed, 07 Aug 2024   Deviance:                       791.73
Time:                        17:49:37   Pearson chi2:                     648.
No. Iterations:                     4   Pseudo R-squ. (CS):            0.01980
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0002      0.046      0.005      0.996      -0.091       0.091
PHLf_value     0.1876      0.040      4.647      0.000       0.108       0.267
==============================================================================
```

Figure 19. GLM with Poisson Distribution Results

The results of comparing the LLF, AIC, and BIC are seen in Table 16. Note that the values of each are much closer to each other when compared to the model used for simulating the PHLf case study data in Table 14. This suggests that each model comparatively fits the data similarly. Providing additional predictor variables may have improved the fit but was intentionally not done to avoid over fitting of the data.

Table 15. Results of Both Models

	GLM w/ Poisson	**Ordinal Logistic Regression**
LLF	-1354	-1283
AIC	2715	2581
BIC	2730	2615

Similar scores for the LLF, AIC, and BIC also suggest that an additional method should be used to test how well the model may simulate the results. Therefore, each model was used to predict the PAS value based on the PHLf score of the case study data and the mean absolute error (MAE) was used as a measure for accuracy of the predictions [73]. The MAE is defined by the following equation:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

Where:

- y_i is the predicted value, in this case the PAS predicted from the statistical model.
- x_i is the actual value, in this case the actual PAS from the case study data.
- n is the number of predictions, in this case 1,064 as those were the values from the case study.

The MAE was calculated for both the GLM with Poisson distribution and the ordered model. The results in Table 16 and Figure 20 illustrate the results for both models with the kernel density estimation displayed to better distinguish between the case study data and the simulated data.

Table 16. Results of Both Models

	GLM w/ Poisson	Ordinal Logistic Regression	Absolute value of difference
MAE	0.988	0.866	0.122

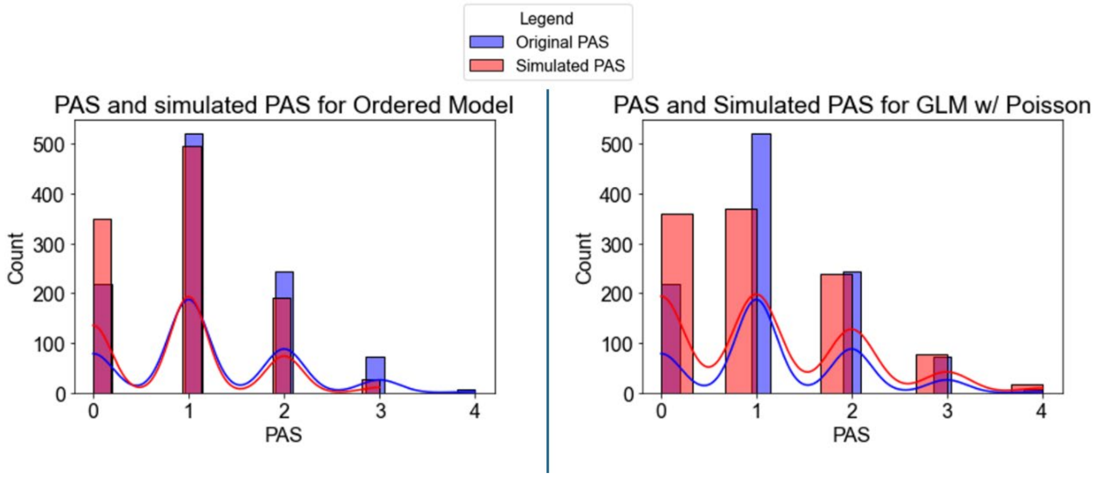

Figure 20. Simulated PAS results for both of the statistical models.

Both Table 16 and Figure 20 show that the ordered model is able to inform the PAS value more than the GLM with Poisson distribution. The GLM with Poison distribution had lower accuracy for a lower PAS, while the ordered model did not produce any incidents with PAS of four, it overall fit the case study data more accurately. Therefore, we can use the ordered model to simulate what the institutional response could have been based on simulated data.

2.Simulated PAS Values for Several Scenarios

This section looks at simulating institutional response against several scenarios. First the ideal scenario where the PHLf score perfectly informs the PAS, then the simulated scenario based on the ordered model from before, a completely random scenario, and an analysis of all three scenarios. From the case study data, it is known that the average number of incidents in a given year were 291 incidents reported to the injury and illness (I&I) program with the lowest being 184 incidents and the maximum being 491 in a given year. For this reason, 290 incidents were simulated using the pSIF model to mimic the number of cases in a single random year.

a. Ideal Scenario: PHLf Perfectly Informs PAS

This scenario is not particularly interesting besides it being the ideal scenario. The objective being that there is a perfect correlation between PHLf and PAS. Which is to say that a one-point increase in PHLf would result in a one-point increase in institutional response. Figure 21 illustrates the scenario, and the following scenarios will follow similar format in illustration.

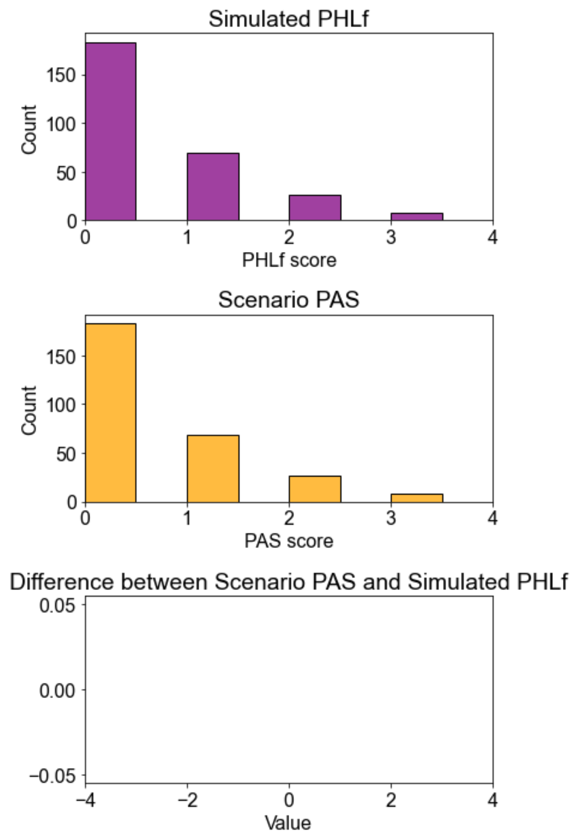

Figure 21. Ideal Scenario of PHLf and PAS

b. Modeled Response Scenario: Statistical Model Informs PAS

In the following scenario, the same PHLf data is used in the statistical modeling of the case study data's PAS values. Figure 22 illustrates the predicted PAS values, and although it seems that the institutional response tends to have a higher response per PHLf score, the difference between each incident PHLf and PAS shows the variance between response. If a PHLf score is two and there is a PAS of three, the institution's response was higher than the PHLf score and it is possible that the response required more resources. This is not necessarily a bad scenario, but if the difference between the PHLf score and PAS is negative, it is possible that the response was not impactful enough for that specific incident.

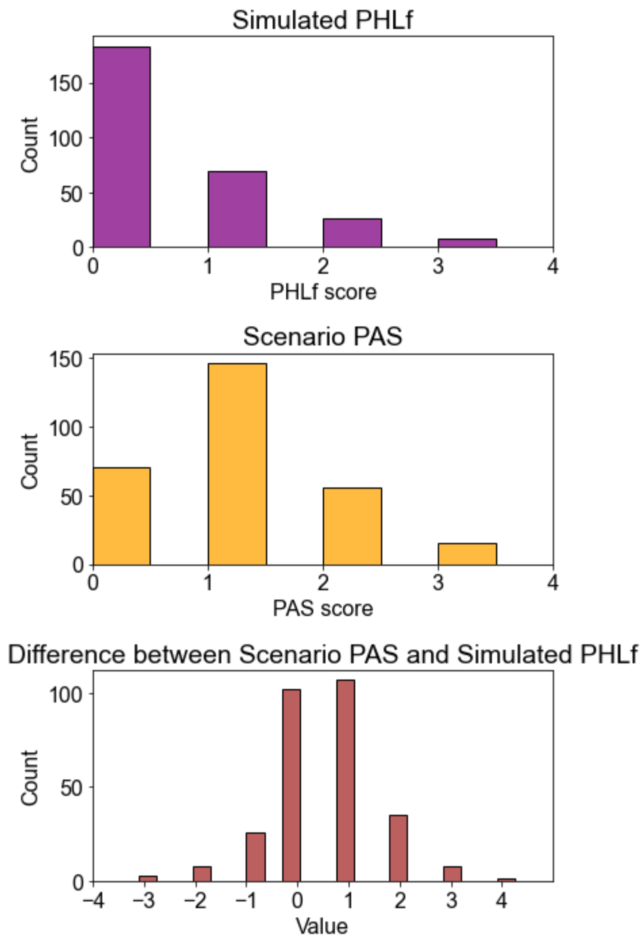

Figure 22. Modeled Scenario of PHLf and PAS

Table 18 shows the number of occurrences where the difference was a certain value for the scenario. In incidents where the difference is three or greater could indicate that the institutional has over responded to an incident that has a low PHLf score. This presents an opportunity to focus resources on those incidents where the PAS score was lower than the PHLf score. Especially since a difference of three would indicate a response that affects several facilities or the whole institution for an incident with a PHLf score of at most one. Inversely, incidents where a difference was lower than negative one present opportunity where the institution should have considered a higher response in an attempt to mitigate the incident. This is especially true for those with score of negative three, as this may suggest an exceptionally low institutional response for an incident that at least had a moderate potential to have been a pSIF. Incidents where the difference is negative ultimately present opportunities to mitigate pSIF and suggest instances where the institutional PAS could have been higher to mitigate future instances of the incident.

Table 17. Difference of PAS and PHLf for Modeled Scenario

Difference	-4	-3	-2	-1	0	1	2	3	4
Number of Occurrences	0	3	8	26	102	107	35	8	1
Percent of Total	0%	1%	2.8%	9.0%	35%	37%	12.1%	2.8%	0.3%

c. Random Response Scenario: Institutional Response is Random

The next scenario looks at the difference between PHLf and PAS if the PAS was completely random. This scenario serves as a control to be able to compare the difference between completely random institutional response and the ideal scenario. A random seed was chosen to enhance the repeatability of this scenario. In this scenario, for each PHLf score a random PAS value between 0 and 4 was assigned. Figure 23 illustrates this scenario in similar fashion to the previous two scenarios.

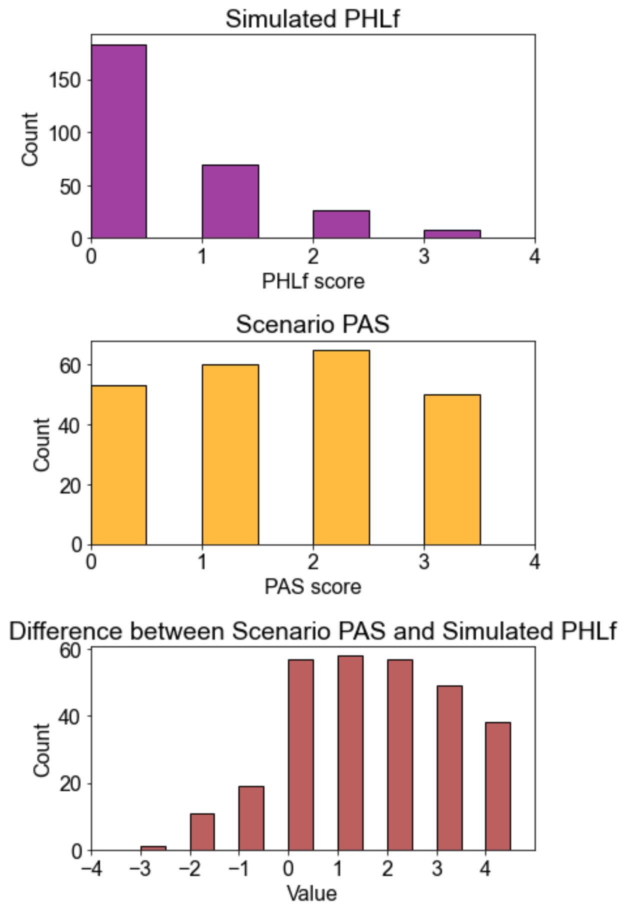

Figure 23. Random Scenario of PHLf and PAS

Table 18 shows the difference between PHLf and PAS for the random scenario. Because the majority of PHLf values are zero, it is an expected result that the majority of incidents in the random scenario led to a higher PAS score. Operationally, this scenario is also not ideal as such extreme responses to incidents with minimal pSIF potential causes a burden on allocated resources. However, in the random response there was a similar number of occurrences where the PAS score was lower than the PHLf. The modeled scenario showed 37 occurrences where the PAS score was below the PHLf whereas the random scenario showed 31 occurrences. Operationally, these occurrences should be minimized in order to mitigate incidents that had a pSIF potential. Next, an examination of the case study in a similar fashion is done to see if any similar inferences can be made.

Table 18. Difference of PAS and PHLf for Modeled Scenario

Difference	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**
Number of Occurrences	0	1	11	19	57	58	57	49	48
Percent of Total	0%	0.3%	3.8%	6.5%	19.7%	20%	19.7%	16.9%	13.1%

d. Case Study Scenario

Figure 24 illustrates the PHLf, PAS, and difference from the case study. The case study data shows that there were no instances where the difference between PAS and PHLf were less than negative two. This shows that none of the incidents with a PHLf score of three had a PAS of zero. Table 20 shows the difference between PAS and PHLf for the case study.

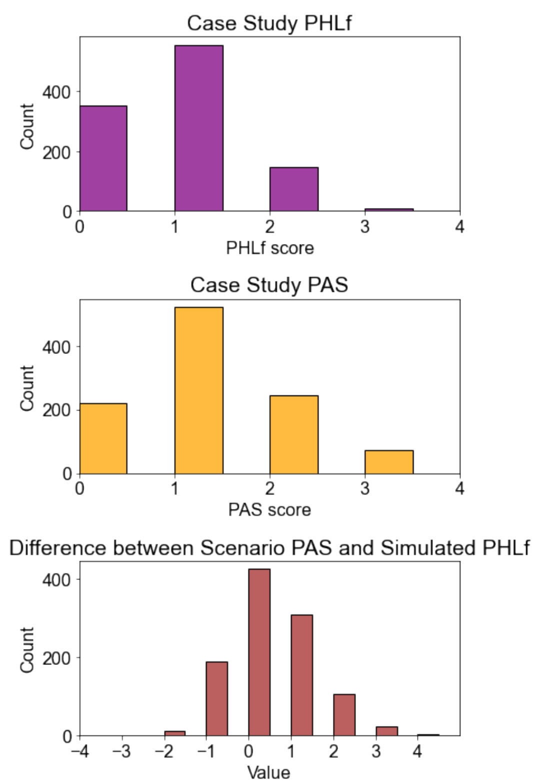

Figure 24. Case Study Scenario of PHLf and PAS

Table 19. Difference of PAS and PHLf for Modeled Scenario

Difference	-4	-3	-2	-1	0	1	2	3	4
Number of Occurrences	0	0	12	188	426	310	105	22	1
Percent of Total	0%	0%	1.1%	17.7%	40.0%	29.1%	9.9%	2.1%	0.1%

F. LIMITATIONS OF PAS AND SCENARIOS

Three separate scenarios were created to model the PAS and the PHLf of those scenarios. The ideal scenario was one in which the PHLf score matches the PAS. Operationally, this scenario concludes that there is perfect correlation between the PHLf of an incident and the PAS. The PAS is a measure of institutional response in the scenario. However, there are some limitations to this approach.

The case study showed no instances of incidents with a PHLf score of four. This is inherently a limitation of the sample that was taken. For this reason, a PHLf distribution modeled after the case study data was used. There are instances where an organization may want to mitigate risks that have a low probability to turn into a SIF incident. This includes occupational hazards that affect the work population. For example, facilities built with asbestos material may provide a minimal risk if there is proper abatement and it would benefit any institution in mitigating those risks before there is an increased potential for the material to deteriorate and become an increased hazard risk. The abatement of this material is typically costly and could affect large work populations.

Additionally, the scenario that is modeled from the case study data also showed instances where the difference between the PAS and PHLf was negative three. Instances where the difference was negative three were not found in the case study data. In fact, the case study data suggests that there were 105 instances where the difference between the PAS and PHLf were two. This suggests that there was a higher response to incidents that had a low potential to cause SIF. Although these situations can lead to mitigation of occupational hazards, it is useful to keep in mind the allocation of resources.

Another limitation of the case study and these scenarios was that the actual severity of the cases was not included. Ideally, actual severity should be reviewed to see if the PAS correlates with the potential severity of the incidents. This may also explain the higher response for some of the incidents from the case study. In instances where an incident causes a SIF to occur, the model should be flexible enough to not penalize the institution for a higher PAS.

Lastly, the PAS is only one aspect of response from an institution and does not give a holistic view of the resources required for proper risk mitigation. Safety management systems are integrated into the complete work process and implementation of a general potential severity framework to other organizations should consider other aspects of risk mitigation beyond the amount of people that are affected by the mitigation.

V. CONCLUSION AND RECOMMENDATIONS

One of the objectives for this thesis was to determine if the current injury and illness (I&I) process at LLNL could be adapted to utilize a potential serious injury or fatality (pSIF) framework. A case study was conducted to assess the feasibility using historical LLNL data. The quality of a sample set of incidents was evaluated to assess whether the current LLNL process captures the necessary data for a pSIF model. Since over 86% of the incidents in a representative sample from 2007 to 2022 had sufficient information in the case data elements, minimal changes would be needed to incorporate a pSIF framework operationally.

The recommendations are to increase efforts in gathering information related to both the root causes of incidents and hazard identification. Among the four criteria reviewed for data quality, these two were found to be the lowest. LLNL has a rigorous work planning process that would be best utilized in hazard identification. Efforts are underway in looking at how to incorporate the work planning control in the I&I process, which would aid in hazards identification.

The case study included the creation of a job-aid for practitioners to utilize when classifying incidents and determining the final potential hurt level (PHLf). Ideally, practitioners would review a set of incidents to calibrate the largely qualitative assessments needed for addressing the diverse range of hazards and controls present in a research and development environment. From the over 1,000 cases reviewed for the case study, only eight cases had a PHLf score of three and none were found to have a score of four. Although this was expected, as serious injury or fatality (SIF) incidents are rare occurrences, this provides a limited resource for performing tracking and trending of these incident types.

Additionally, the case study also aimed to evaluate institutional responses by qualitatively reviewing the impact of corrective actions recorded within the I&I process. A review of incidents with a PHLf score of two identified 142 cases, accounting for approximately 13% of the total incidents analyzed. The top three accident type categories—noise exposure, overexertion, and repetitive motion—comprised 62% of the cases with a PHLf score of two. This finding highlights opportunities for improvement and mitigation through engineered controls, such as noise suppression, and administrative controls, like adjusting the frequency of work tasks to prevent repetitive motion injuries.

It is imperative to keep in mind that research institutions are constantly repurposing equipment and lab space to meet changes in work scope. Preventative actions in the form of work control are also essential for mitigating long-term organizational risk. These considerations should inform institutional responses to incidents, as significant policy changes can have lasting implications for work done in an organization.

The institutional response, noted as the Preventative Action Score (PAS), to injuries suggests that the higher the potential severity of the incident, the greater the response, even without the implementation of the framework. However, the distribution of PAS values was not consistently applied across the representative sample. This suggests that a pSIF framework can improve organizational risk management by providing a more consistent approach to incident response. Appendix C also highlights how a pSIF framework could be generalized for other frameworks, offering organizations a holistic view of their risk management.

The computational modeling of incidents and scenarios provides deeper insights into improvements that a pSIF framework can facilitate as well as model scenarios not scene in the representative sample. Three scenarios were modeled based on the average number of injury and illness (I&I) incidents may occur in a given year. The computational model found that scenarios where the difference between the PAS and PHLf was less than negative one should be revisited to ensure that a proper institutional response is provided. The model also provides an opportunity to further refine the PAS and PHLf methodologies and to forecast institutional responses based on PHLf distributions. The outcome of the computational model was compared to the results of the case study and found that the modeled scenario underrepresented institutional response.

The case study data showed that the PAS was typically higher than the PHLf value for an incident. This meant that the institutional response to incidents typically used more resources than initially expected. Although scenarios where the PAS exceeds the PHLf are not necessarily adverse conditions, it may indicate instances where a less resource intensive method of mitigation should be considered due to the probability of a pSIF occurring. This approach may free up resources to address incidents with a PHLf score of three or four, indicating a higher likelihood of causing a SIF.

A recommendation from the results is for organizations to use the classification of incidents to inform resource allocation towards those with the greatest risk. When developing metrics for pSIF implementation, one should consider the difference between PAS and PHLf to identify opportunities for improving risk management. The modeled scenario found that there were 31 instances where the PAS was less than the PHLf, indicating a possible suboptimal response from the institution.

Another recommendation is to view the pSIF framework as an integrated approach to risk management that enhances organizational impact. This aids the redistribution of resources across management systems, emphasizing integrated improvement. With appropriate modifications, the pSIF framework can be implemented to add value to integrated management systems. This approach would focus on a holistic perspective to risk mitigation.

Modifications to the PAS are necessary to ensure alignment with organizational resource allocation. Any significant changes within an organization present inherent risk and implementing this framework is no exception. It requires an implementation plan, additional training, and the comprehensive tracking and trending of incidents to inform those with the highest organizational risk. Insufficient attention to lower severity incidents may increase their frequency, potentially elevating overall organizational risk. Additionally, there is a need for consensus on the thresholds for high severity incidents and to agree on the use of incident classification to inform response and corrective actions.

For this reason, the scope of the thesis included the computational modeling of relevant case studies, as well as modeling of the relationship between PHLf score and PAS. The PAS is a generalized method for classifying institutional response. However, the computational models developed were simplistic, which highlights opportunities for greater improvement.

Further work is needed to validate the efficacy of the framework. Specifically, addressing the limitation of the case study will require testing the model against various DOE OPEX incidents [50]. This approach could prove an effective way to evaluate the operational pSIF model against known incidents to ensure the thresholds for each category are appropriate. This incident database includes environmental and security incidents, providing an opportunity to test a generalized framework beyond safety as described in Appendix C. The ability to test against larger datasets is crucial, especially since higher severity incidents are rare occurrences across various management systems; this was a clear limitation of the case study sample set.

Lastly, implementation of the framework and tracking of key performance indicators are required to ensure maximum efficacy. This requires a clear understanding of current institutional response to risk. Further research includes retrospective analysis of incidents to create example scenarios for framework implementation. The scope of national research laboratories are ever changing and the framework outlined can provide an additional toolset in comprehensive risk management.

APPENDIX A. EXAMPLE SCENARIO

The following is an example incident to show case how a proposed pSIF model would work in operation. The scenario is not intended to represent any real events and any resemblance to an actual incident or personal experience is purely coincidental. The scenario is adapted from the case study performed but tailored more for a nuclear operations research and development environment [15].

A. SCENARIO

"A laboratory technician working with inorganic compounds in a glovebox noticed a strong odor during routine operations. The technician noticed a strong odor and mucous membrane irritation during the operations, which regularly produces hydrogen sulfide. The operations regularly produce hydrogen sulfide. The technician alerted others in the area and warned them to leave immediately. A stop work was initiated after informing the supervisor. The technician is transported to the on-site clinic and is tested for chemical exposure. The safety team was notified to ensure safe reentry and to inspect the glovebox. An industrial hygienist reviewed the work and assessed that the quantities of material within the glovebox operations could have led to respiratory depression. After inspection of the glovebox, it is noted that there was an improper seal from the gaskets that had deteriorated over time, leading to the incident. Consequently, the gloveboxes in the facility were inspected for leaks."

B. DETERMINING FRAMEWORK VARIABLES

Using the job-aid in Figure 8, the first step is to determine what could have been the potential hurt level worst-case (PHLwc) for the incident. A reasonable worst-case includes a review of the quantity of hazardous materials and types of hazards. Given an example of quantities where inadvertent chemical exposure could have led to some type of respiratory depression or damage to the central nervous system may be possible. If the main hazards from within the glovebox are sparking hazards and the chemical hazards are secondary, this would also be reviewed at this stage. In this example an industrial hygienist that is comfortable and knowledgeable of the work occurring in the laboratory assessed the hazards and assigns a PHLwc score of 3 (severe hurt) based on Table 5.

Next, the primary variable (V) is identified. This is where a review of possible variables that would have contributed to the PHLwc score is considered. Possible contributing variables include the amount of material used, body placement, training, or ventilation system failure. The investigator concludes that the material used was the primary variable, with a likelihood of the material being used leading to the worst-case being around 50% of the time. This leads to a primary variable of "sometimes."

The second variable required is the controls effectiveness (C), a qualitative measure of hazard mitigation through controls. It is a holistic approach to all controls in place at time of the incident, even if they were ineffective. In this case, this would include items such as administrative controls like procedures, training, glovebox specifications, alarms, and personal protective equipment. In this case, the glovebox as an engineered control was ineffective. Using the V/C score matrix in Table 6, the value is determined to be three.

With the V/C score and PHLwc determined, Table 7 is used to determine the final potential hurt level (PHLf). The PHLf was then determined to be three, indicating that there was a high potential for a serious injury or fatality and would warrant an elevated response from the institution.

C. INSTITUTIONAL RESPONSE

For the framework to be effective, it is not only pertinent to evaluate incidents for their PHLf score, but to respond effectively to the mitigation of such incidents with higher scores. Because the example incident leads to a PHLf determination of three, it is recommended that there is an institutional response of three or more in risk mitigation.

The response would most likely include a root cause to determine mitigation. In the example, it is possible that there were modifications to the glovebox or perhaps the frequency of testing glovebox for those operations needs to be reevaluated. In the incident report, it is stated that the gloveboxes in the facility were inspected, which would be classified as a PAS of three based on Table 8.

If the operations occurring within the glove box lead to the failure in engineered control and if warranted could be escalated to a score of 4. This response may look like changes to policy affecting current and future operations. The proactive risk mitigation approach considers the precursors that led to the event.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. ALTERNATE STATISTICAL METHODS TESTED

The following appendix explains the other two methods of statistical modeling that engaged in testing throughout the thesis, along with the results. The evaluation and testing of these statistical models were important in informing the most ideal statistical modeling method for the computational model.

A. GENERALIZED LINEAR MODEL WITH A POISSON DISTRIBUTION

GLM with a Poisson distribution is used where the response to a predictor variable is the count of incidents that occurred within a certain amount of time. Poisson distributions can be used to calculate the number of rare events, such as radioactive decay, uncertainty analysis, and safety. Understanding these incidents made it a clear choice to attempt to model the number of events in each category over a period of time, perfect for simulating decades worth of data. However, Poisson distributions have an assumption that the mean and the variance of the data is equal. For the case study, this would mean that if over the course of five years the average number of incidents with a PHLf score of four was five, there would be a variance of five over the same period. Because fatalities at LLNL have been so rare and the case study found no incidents with a PHLf score of four, it is possible that under dispersion may occur. The model is defined with the following equation:

$$log\left(E\left[\, Y_{PHLf} \mid X_{PHLwc}, X_V, X_C, X_{VC} \,\right]\right) =$$

$$\beta_{PHLwc}X_{PHLwc} + \beta_V X_V + \beta_C X_C + \beta_{VC} X_{VC} + \beta_i \qquad [46], [47].$$

Where:

- $log(E[\, \mathrm{Y_{PHLf}} \mid X_{\mathrm{PHLwc}}, X_{\mathrm{V}}, X_{\mathrm{C}}, X_{\mathrm{VC}} \,])$ is the log of the expected count of the responses given the predictor variables.
- $\beta_{\mathrm{PHLwc}}X_{\mathrm{PHLwc}} + \beta_{\mathrm{V}}X_{\mathrm{V}} + \beta_{\mathrm{C}}X_{\mathrm{C}} + \beta_{\mathrm{VC}}X_{\mathrm{VC}}$ represents the product of the coefficients, β-value, and the predictor value, X, that defines the threshold for a one-unit change for the PHLf score.

- β_{i} is the constant that defines when the PHLf score, $\mathrm{Y_{PHLf}}$, is zero.

Similarly to the ordinal logistic regression model, the generalized linear model with a Poisson distribution uses MLE to find the best fit. An explanation of how MLE is used and how it is different from the method of ordinary least squares in linear regression is explained in the previous section, Section A. 1. Ordinal Logistic Regression Model. An important distinction of generalized linear models is that it allows the response variable to have distributions different from typical normal distributions [46]. The fact that typically injury and illness incidents occur at a time independently of other incidents is also an assumption of the Poisson distribution. Poisson is also a discrete probability distribution, which means in this case there are no instances of the negative or partial incident. An incident either occurred or did not. The Poisson distribution of the predictor variables is defined by the following equation:

$$p(X_i; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad [46], [62].$$

Where:

- $\mathrm{X_i}$ is the predictor variables in the model such as $X_{\mathrm{PHLwc}}, X_{\mathrm{V}}, X_{\mathrm{C}}, X_{\mathrm{VC}}$.
- λ represents the shape parameter indicating the average number of events and the variance of that value.
- x is a non-negative integer and the count of PHLf scores in each category.

The Poisson equation above explains the probability that the random variables from above take on a certain value. The shape parameter, λ, is useful in understanding how the data is skewed and has been used to explain injuries in military operations [63] and has been cited as being dependable for analysis of traffic accidents [64]. Making it an ideal distribution to simulate incidents.

B. ORDINAL LOGISTIC REGRESSION MODEL

The ordinal logistic regression technique is common for data sets where the data is categorical and follows an order, such as the Likert scales type questions seen on satisfaction surveys. Another important assumption about this method is that the ranking of each level does not mean that the intervals between the ranks are equal. For example, a PHLwc of one may indicate a broken arm, but a PHLwc of thee may indicate an amputation. An amputation is not compared by proportions of broken arms. The regression method uses the predictor variables V, C, and V/C, to predict the probability that the PHLf falls into a specific category. The model is defined with the following equation:

$$log\left(\frac{P(Y_{PHLf} \leq j_{PHLf})}{1-P(Y_{PHLf} \leq j_{PHLf})}\right) = \beta_{PHLwc}X_{PHLwc} + \beta_V X_V + \beta_C X_C + \beta_{VC}X_{VC} + \alpha_j \text{ [47]}$$

Where:

- $\log\left(\frac{P(\mathrm{Y_{PHLf}} \leq \mathrm{j_{PHLf}})}{1-P(\mathrm{Y_{PHLf}} \leq \mathrm{j_{PHLf}})}\right)$ is the log of the probability that the PHLf score, $\mathrm{Y_{PHLf}}$, is a score of $\mathrm{j_{PHLf}}$ or lower and is also known as the logit function [47].

- $\beta_{\mathrm{PHLwc}}X_{\mathrm{PHLwc}} + \beta_{\mathrm{V}}X_{\mathrm{V}} + \beta_{\mathrm{C}}X_{\mathrm{C}} + \beta_{\mathrm{VC}}X_{\mathrm{VC}}$ represents the product of the coefficients, β-value, and the predictor value, X, which defines the threshold for a one-unit change, and the predictor variables, X.

- α_j is a constant that will be used to define the threshold between categorical scores of PHLf.

Similar to how linear regression uses a method of ordinary least squares, logistic regression uses the maximum likelihood estimation to find the best fit. The model works by adjusting the coefficients until the likelihood that the series of predictor variables results in the target outcome is maximized. There are several methods that can be used to iteratively optimize the algorithm to solve the equation. Among these, the BFGS [49] method was used and is commonly used for optimization problems. BFGS optimization works by approximating the second-order partial derivatives of the function to get a Hessian matrix. The matrix is refined at each iteration until the local extrema of the function is found. The outcome of the regression model is therefore the probabilities of the output for each value of the predictor variables [47]. However, another regression model was explored to simulate the data.

Now that there are two models, they can be leveraged to create simulations of the case study data. Because the case study is based on a random sample of over ten years, it may be a good indicator of what the distribution of injury and illness cases look like and can be used to synthesize data of various distributions. For the thesis, the goal is also to synthesize the case study data to model institutional response for the development of useful metrics in the case of framework implementation. Creating these metrics is important in tracking framework influence in the organization [56]. The next section will go into detail on simulating the case study data for modeling implementation.

The models provide the statistical parameters to provide the distribution of PHLf scores. The results showed that PHLf, regardless of model, is mostly influenced by the PHLwc and the V/C value. Which is to say that there is no statistical correlation between the V and C score. For this reason, the PHLf and V/C value will be used to generate the synthetic data in question. Figure 25 shows the code that was used to simulate the data from the ordinal logistic regression model.

Simulating results using the OrderedModel

```
In [48]:
def simulate_data(sample_size, beta_PHLwc, beta_VC_value, thresholds, mean_PHLwc, std_PHLwc, mean_VC_value, std_VC_value):
    # Using the mean and standard deviation from the model
    PHLwc_values = np.random.normal(loc=mean_PHLwc, scale=std_PHLwc, size=sample_size)
    VC_value_values = np.random.normal(loc=mean_VC_value, scale=std_VC_value, size=sample_size)

    # Calculate the linear predictor values
    linear_predictor = beta_PHLwc * PHLwc_values + beta_VC_value * VC_value_values

    # Calculate the cumulative probabilities for each threshold
    cumulative_probs = np.array([expit(linear_predictor - threshold) for threshold in thresholds])

    # Calculate the probabilities for each category
    probs = np.diff(np.vstack([np.zeros(linear_predictor.shape), cumulative_probs, np.ones(linear_predictor.shape)]), axis=0)
    probs = np.clip(probs, 0, 1)
    probs = probs / probs.sum(axis=0)

    # Generate values based on the probabilities
    simulated_PHLf_values = np.array([np.random.choice(len(thresholds) + 1, p=prob) for prob in probs.T])

    # Apply constraints to ensure values do not exceed framework values
    PHLwc_values = np.round(PHLwc_values).astype(int)
    VC_value_values = np.round(VC_value_values).astype(int)
    return pd.DataFrame({
        'PHLwc': PHLwc_values,
        'VC_value': VC_value_values,
        'PHLf_value': simulated_PHLf_values
    })
```

Figure 25. Code Snippet to Generate Synthetic Data for the Ordinal Logistic Regression Model

The function takes the various β-value of the two coefficients that had the most influence in determining the PHLf score, its thresholds, the mean, and standard deviation of the coefficients. The sample size variable is used to know the amount of simulated values, or PHLf scores, to generate. The thresholds are used to categorize the generated values. An ordinal logistic regression does not have an assumption of the distribution of the values. Because a Poisson distribution was not assumed, a normal distribution was used. The normal distribution is defined by the means and standard deviations from the model in order to generate the characteristics of the predictor variables in the model.

The linear predictor is used to create a weighted sum of generated predictor values. In this manner, the function computes the relationship between these values and the outcome. However, these linear predictor values must be transformed into a cumulative probability to be deterministic. This is done using the logistic sigmoid function, defined as $\sigma(x) = \frac{1}{1+e^x}$, used as SciPy's Expit function [69]. Using the probabilities for each category, the code then ensures that the probability will be within zero and one. This is to ensure that the probability is possible. Then the values are generated based on the probabilities and the PHLwc scores and V/C scores are converted from continuous values to discrete integers. The simulated values were then saved in a Pandas data frame and the results for simulated values of 900, 1,000, and 1,100 cases can be seen in Figure 26. An important thing to note is that a set value for the random seed was used to get repeatable results.

An important feature to note is that although the PHLwc and V/C distribution is similar to that of the original case study, the distribution in PHLf is biased on the lower and upper bounds of the values from the case study. This is clearly problematic in terms of simulating the PHLf distribution for modeling changes in distribution. For this reason, an attempt was made to use the GLM with Poisson distribution to see if a more similar distribution could be simulated. It is possible that there were several issues with the model. One of which was that the ordinal logistic regression model could be overfitting the data. This also explains why there is no coefficient for where the PHLf value was four. However, because the Poisson distribution is specifically for counts of events, it may be able to better simulate various distributions of injury and illness data.

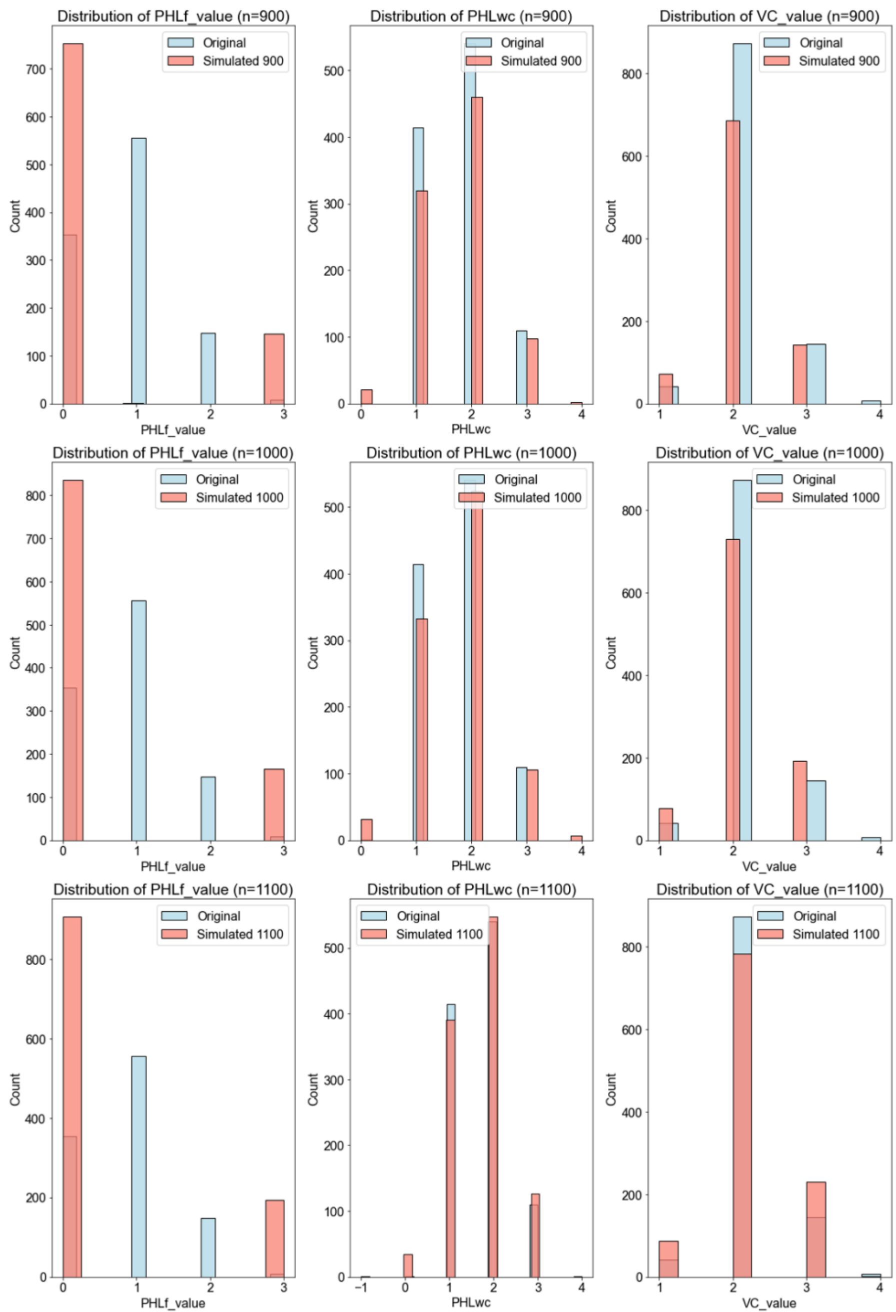

Figure 26. Results of Ordered Model PHLf Simulation

The results of the GLM with Poisson distribution show that there is a positive correlation between PHLf and both the PHLwc and V/C score. Similarly to the ordinal logistic regression model, it seemed that the values of V and C showed less of a statistically significant correlation. Although the LLF, AIC, and BIC of both models were significantly different, it is possible that the Poisson distribution may be more robust when simulating the data. Figure 24 shows the code that was used to generate simulated results. Like the function that was created to simulate data from the ordinal logistic regression model, the GLM with Poisson distribution uses the β-values, mean, and standard deviation of the predictor variables.

Created simulated results based off the Poisson GLM model

Next the distribution results are used as a way to create simulated data.

```
In [11]:
# Function that generates simulated sample sizes based on the distribution of the model or given beta values
def simulate_data(sample_size, beta_PHLwc, beta_VC_value, mean_PHLwc, std_PHLwc, mean_VC_value, std_VC_value):
    # Using the mean and standard deviation from the model
    PHLwc_values = np.random.normal(loc=mean_PHLwc, scale=std_PHLwc, size=sample_size)
    VC_value_values = np.random.normal(loc=mean_VC_value, scale=std_VC_value, size=sample_size)

    # Calculate the lambda for the predictor values
    lambda_values = np.exp(intercept + beta_PHLwc * PHLwc_values + beta_VC_value * VC_value_values)

    # Generate Poisson-distributed outcomes
    simulated_PHLf_values = np.random.poisson(lam=lambda_values)

#     # Apply constraints to ensure values do not exceed framework values
#     VC_value_values = np.clip(VC_value_values,0,4)
#     PHLwc_values = np.clip(PHLwc_values,0,4)

    PHLwc_values = np.round(PHLwc_values).astype(int)
    VC_value_values = np.round(VC_value_values).astype(int)
    simulated_PHLf_values = np.round(simulated_PHLf_values).astype(int)
    #Apply constraints only to simulated PHLf values
    simulated_PHLf_values = np.clip(simulated_PHLf_values, 0, 4)


    return pd.DataFrame({
        'PHLwc': PHLwc_values,
        'VC_value': VC_value_values,
        'PHLf_value': simulated_PHLf_values
    })
```

Figure 27. Code Snippet to Generate Synthetic Data for the Ordinal Logistic Regression Model

However, an important distinction is that the GLM with Poisson distribution generates a λ-value which defines the shape parameter of the Poisson-distributed outcomes. An explanation of the equation $\lambda_i = \exp\left(\beta_0 + \beta_{PHLwc} X_{PHLwc,i} + \beta_V X_{V,i} + \beta_C X_{C,i} + \beta_{VC} X_{VC,i}\right)$ is derived in Appendix A. Section A. The shape parameter takes the place of the use of the thresholds of the ordinal logistic regression model and the function then proceeds to generate the values based on the discrete probability. Additionally, because the shape parameter is exponentiated, the lambda value is always positive. Not only is this a requirement of the Poisson distribution, but this also reflects an important assumption that there are no such thing as negative incidents. The values of PHLwc and V/C score at this point are floats because they are generated randomly based on the mean and standard deviation of the original model. For this reason, they are binned into their nearest integer and the simulated PHLf values are limited to only produce cases with a value between zero and four to meet the constraints of the framework. Again, the random values are generated using a known seed for repeatability. The simulated values were then saved in a Pandas data frame and the results for simulated values of 900, 1,000, and 1,100 incidents can be seen in Figure 28.

The results of the GLM with Poisson distribution were much more favorable than the simulated values generated from the ordinal logistic regression model. Both in terms of fitting against the case study data and in terms of the ability to adjust parameters to simulate different parameters. For this reason, the GLM with Poisson distribution was chosen and should be considered when simulating various distributions of the data. Next the focus will be on simulating institutional response to generate metrics for use in an operational setting.

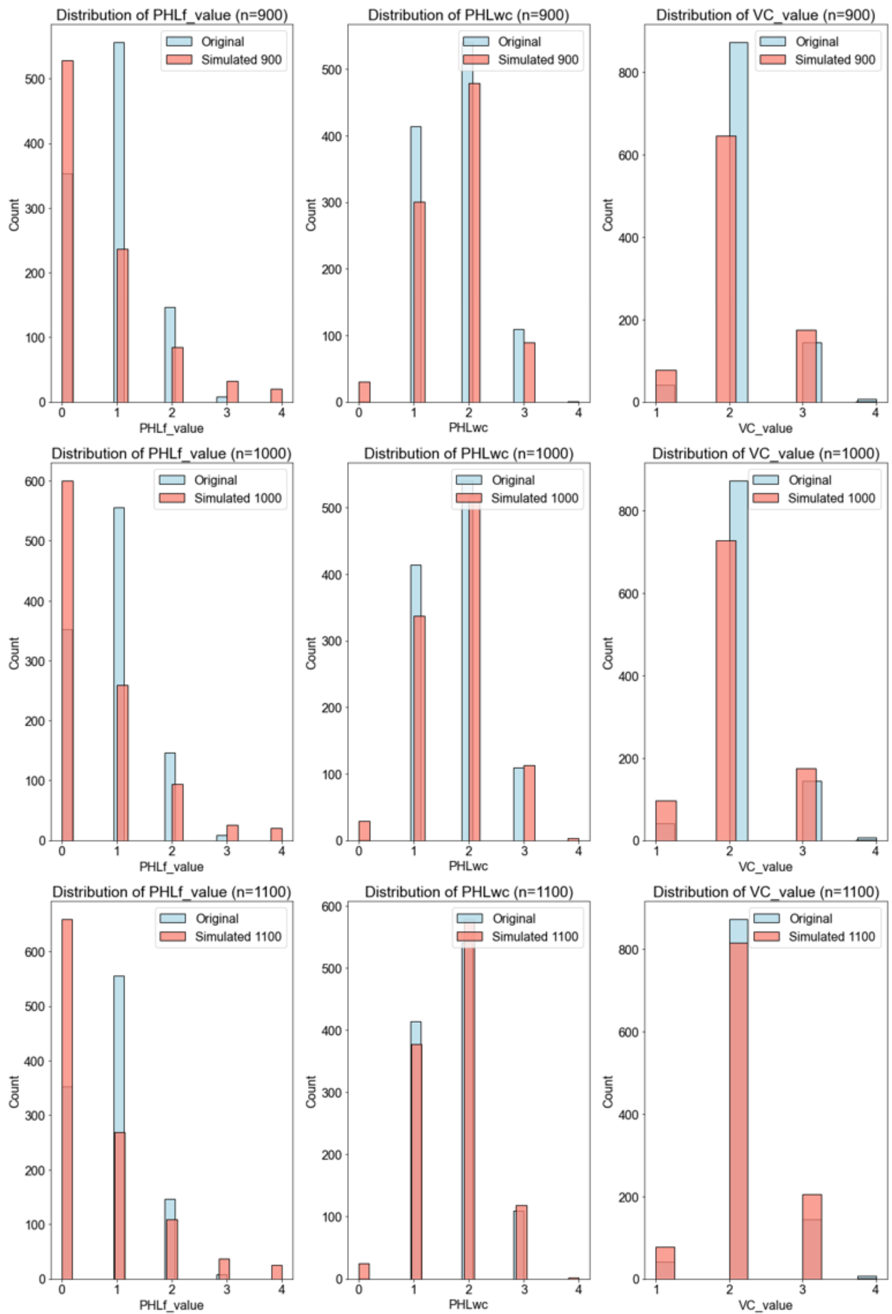

Figure 28. Results of the GLM with Poisson Distribution PHLf Simulation

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. MANAGEMENT SYSTEMS

The Livermore site has environmental and occupational health and safety management systems in place to provide support for protecting its workers, the public, and environmental stewardship [21]. It is imperative that the proposed model live within this management system to be effective. It achieves this through an integrated management system approach that considers the International Organization for Standardization's (ISO) 14001: Environmental Management Systems [22], and ISO 45001: Occupational Health and Safety Management System (OHSMS) [23]. Both ISO standards require and emphasize continuous improvement, evaluation, and risk mitigation to minimize adverse environmental and occupational health and safety risks.

The framework of the case study is meant to demonstrate how to encompass a continuous improvement philosophy in a way that can mitigate high-severity adverse incidents across management systems. A look at the framework and it's fit for purpose against other management systems is discussed in later chapters. The overlap between ISO standards also provides additional structure for the framework to be applicable outside of environment and health. Such conceptual overlap may benefit concerns in quality or security management systems.

A. IMPLICATIONS OF THE CASE STUDY BEYOND HEALTH AND SAFETY

Frameworks that are system agnostic provide a flexible, scalable, and general solution that can be adopted across an organization. For instance, LLNL integrates the ISO 45001 [23] and ISO 14001 [22] standards to simultaneously manage occupational health and safety alongside environmental management. The integration raises insights on how lessons learned from a pSIF framework may be applied beyond occupational health and safety. For this reason, the following sections highlight possible implications across various management systems.

B. IMPLICATIONS FOR ENVIRONMENTAL MANAGEMENT

LLNL follows ISO 14001 [22] for guidance on managing its environmental management system. One of the guiding principles of any environmental management system is its commitment to continuous improvement. Especially doing so in manner that mitigates any negative impact to the environment. DOE national research laboratories are required to follow strict federal, state, and local regulations ranging from executive orders to voluntary obligations. This is all in part of ensuring federal agencies consider the potential environmental impacts in accordance with the National Environmental Policy Act.

In order to correlate a pSIF framework to another type of system, an understanding of what constitutes a high severity incident for that system must be understood.

Environmental incidents can impact organizations in several ways. This includes monetary fines and penalties for non-compliance. Unlike private entities, national research laboratories cannot recuperate losses due to environmental non-compliance. Severe environmental incidents may lead to a degradation in public perception and community relations. As well as leading to increased scrutiny from federal oversight.

For example, Lawrence Livermore National Laboratory is surrounded by a lively community and several wineries. An incident related to the release of toxic material into the water could create irreparable damage to the laboratory's reputation. Understanding the implications of various incident types can facilitate a crosswalk between occupational health and safety and environmental management, as illustrated in Figure 29.

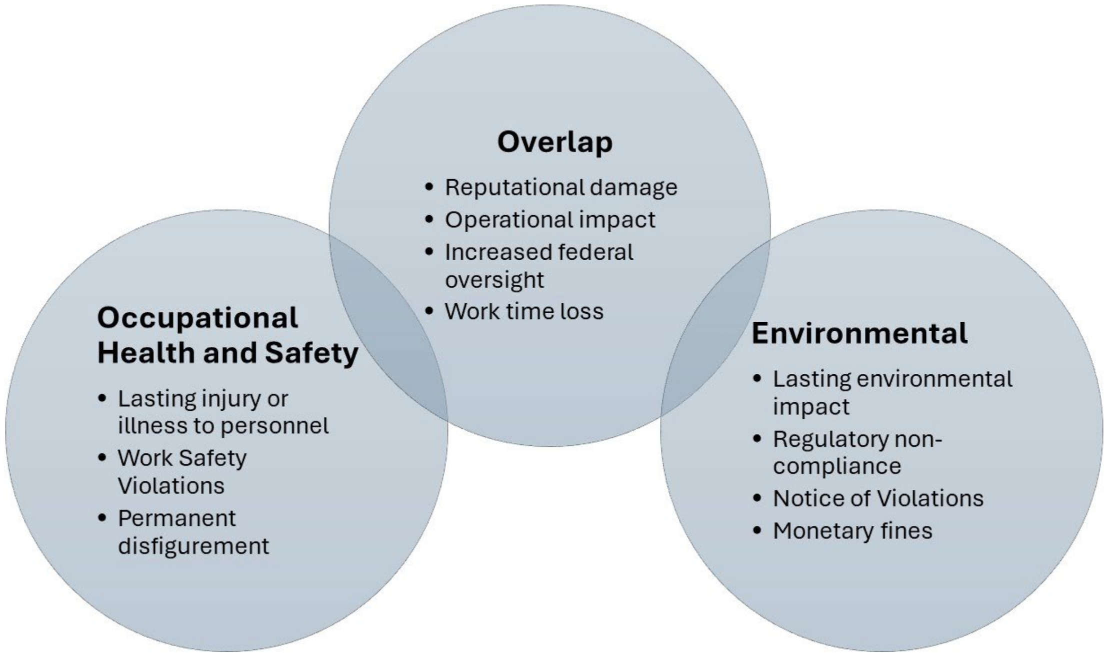

Figure 29. Incident Overlap in Occupational Health and Safety and Environmental

The framework can then be used to evaluate incidents based on overlapping repercussions, while still recognizing some of the unique implications of environmental incidents and their varying severity. A key difference between injury and illness incidents and environmental incidents is that while many minor incidents may not increase the severity of impact to an individual, the same cannot be said of environmental incidents. For example, the release of smaller quantities of hazardous materials to the environment can eventually lead to larger environmental impact.

In contrast, multiple minor injuries would not necessarily lead to a severe outcome. Although multiple strains on muscle may cause irreparable damage it is unlikely that strains would lead to amputation of a limb, except under extreme circumstance. The pSIF model also assumes that a percentage of near misses could contribute to more significant events. This has not been verified for smaller environmental impacts leading to more significant environmental impacts and may require a similar case study to be done. Although at least one study focused on China's chemical industry looked at a general safety triangle in terms of occupational health and safety, environmental, and even quality [72]. This may imply that the overlap and framework may apply to other management systems.

C. IMPLICATIONS FOR QUALITY MANAGEMENT

In addition to ISO 14001 [22] and ISO 45001 [23], LLNL also adheres to ISO 9001 for guidance on managing its Quality Management System (QMS). Similarly to the other management system standards, ISO 9001 emphasizes the principle of continuous improvement. A QMS is responsible for having a system in place to documents processes, procedures, and responsibilities necessary to achieve operational objectives. Given LLNL's mission to support national security research; ranging from stockpile stewardship to fundamental science discovery – the implementation of a QMS can streamline operations, reduce costs, and drive long-term success in meeting contractual agreements.

Understanding the implications of incidents regarding quality can enhance how a pSIF framework impacts QMS and mission objectives. A common consequence of adverse quality incidents is the Cost of Poor Quality (COPQ). COPQ refers to the cost associated when processes are inefficient or produce subpar results [53]. From a mission perspective, catastrophic failures can impact product and service quality leading to far-reaching impacts across the complex. This is especially important as LLNL is the only DOE national research facility that relies on other facilities within the complex to meet research objectives. Such failures may also result in occupational health and safety incidents as well as not meeting mission or contractual obligations.

National research laboratories, such as LLNL, have strong brand recognition that could influence federal and public perception institutional performance. As illustrated in Figure 30, there is an overlap between negative outcomes related to occupational health and safety and quality.

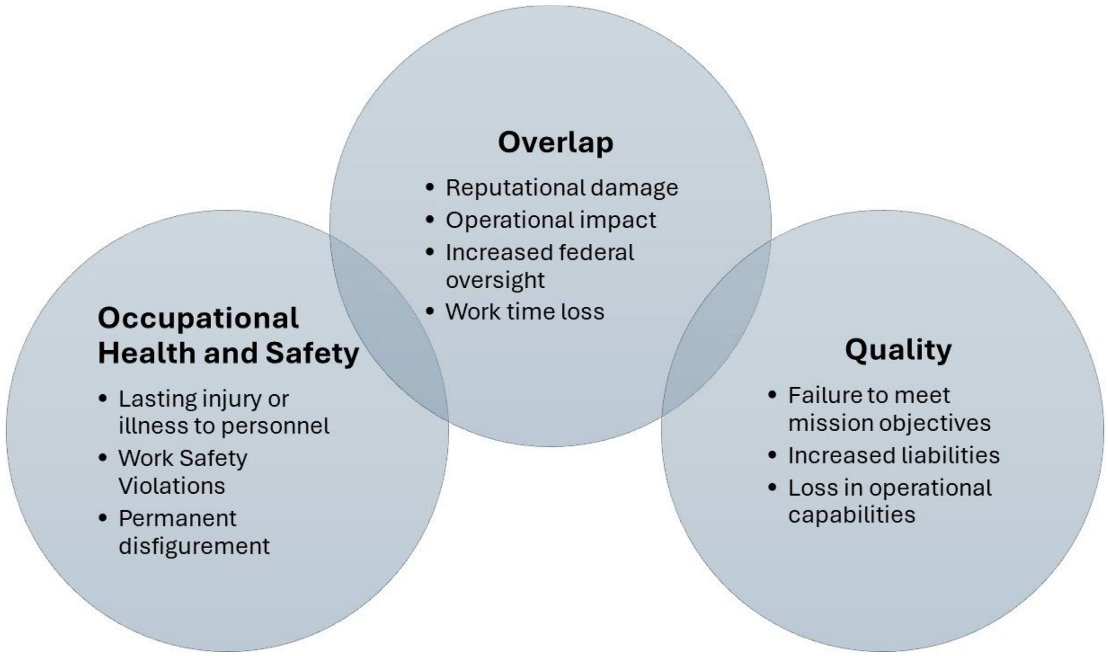

Figure 30. Incident Overlap in Occupational Health and Safety and Quality

Quality management encompasses broad operational processes and for this reason, it is crucial to understand how quality affects both Occupational Health and Safety Management Systems (OHSMS) and Environmental Management Systems (EMS). For example, in review of an incident that lead to an inadvertent dose in radiation could stem from a failure in the management of the quality of a process that lead to inadequate engineered control to prevent it. Thus, systems quality has far-reaching relevancy across all management systems, not just in the production of widgets.

D. IMPLICATIONS FOR SECURITY

Lastly, there was interest in seeing if the pSIF framework could apply to security. A system typically not associated with OHSMS and EMS. Although security is a broad topic that can encompass information security, physical security, and national security. Understanding the implications of high severity incidents in a comprehensive manner can facilitate the implementation of a pSIF framework across these areas. This includes the overlap as a consequence of the incidents from each system.

Security management systems typically integrate people, processes, hardware, and software to apply a risk management strategy to mitigate incidents with a negative outcome. The DOE considers quality, environmental, health and safety, and security in its operational excellence lessons learned [50]. The interconnectedness in Figure 31 illustrates the overlapping consequences between these systems. Security has similarities to OHSMS and EMS in that both management systems must mitigate operational risk, perform incident response, and ensure compliance with federal, state, and local regulations. For these reasons, a pSIF framework may have success in being adapted to various security management frameworks. However, in this instance, security is a general field of study, where operationally, security may encompass physical security, cyber security, and other aspects that each have specific ways to gauge organizational risk.

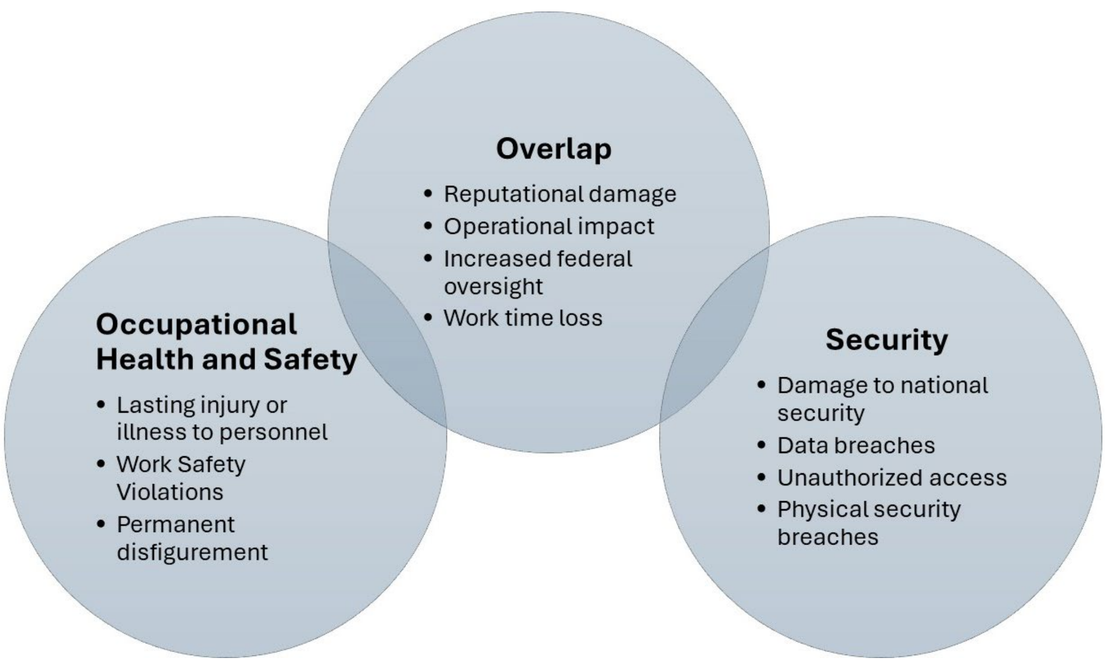

Figure 31. Incident Overlap in Occupational Health and Safety and Broad Security Management

E. OVERARCHING SIMILARITIES IN MANAGEMENT SYSTEMS

By examining the lasting implications of high-severity incidents across different management systems, LLNL can enhance system integration and performance. The

International Organization for Standardization (ISO) released a handbook in 2018 on the integrated use of management system standards, [54] which influenced the expansion of the pSIF framework to encompass multiple management systems. LLNL has already started efforts in creating an integrated management system that combines OHSMS and EMS. There has also been institutional interest in understanding how the interconnectedness of these management systems can enhance operational excellence.

For clarity, Figure 32 synthesizes Figures 13–15 and illustrates overlap in incident outcome between the management systems discussed in Subsections 1–3. Effective post-incident classification based on their overall severity can aid in continuous improvement of these systems.

Lastly, integrated management systems present opportunities for increased efficiency within organizations. For example, a nuclear power plant in East China sighted that implementing an integrated management system assisted in passing multiple certifications and cited reduction in bureaucracies, management reviews, and corrective and preventative actions [55]. The pSIF framework aims in improving corrective and preventative actions, and its adoption across management systems may shift focus from isolated deficiencies to comprehensive system improvements. Siloed management systems can therefore have a holistic approach to risk management.

Figure 32. Overlap Between Several Management Systems and Incident Severity

The concept of the "worst-case scenario" that institutes the highest level of hurt level can then be reframed by the "highest level of incident severity," dependent on the management system involved. Table 20 presents various examples of proposed incident severity level definitions, while maintaining the structure of the original pSIF frameworks used in the LLNL case study [15]. Additionally, Figure 33 shows a modified job aid for a general approach to a potential severity incident framework.

While a generalized approach to the potential severity framework brings a holistic approach to risk mitigation in an organization, it presents opportunities for confusion in terms of institutional response. But it is still important to recognize that incidents in other management systems may have contributing factors that could escalate their potential severity. This is an underlying assumption of the pSIF framework. For example, if an organization breaks up its management systems across the institution to various directorates or departments, the roles and responsibilities may create an issue of fragmented responsibilities. This has the potential to lead to diminished accountability in incident response [52].

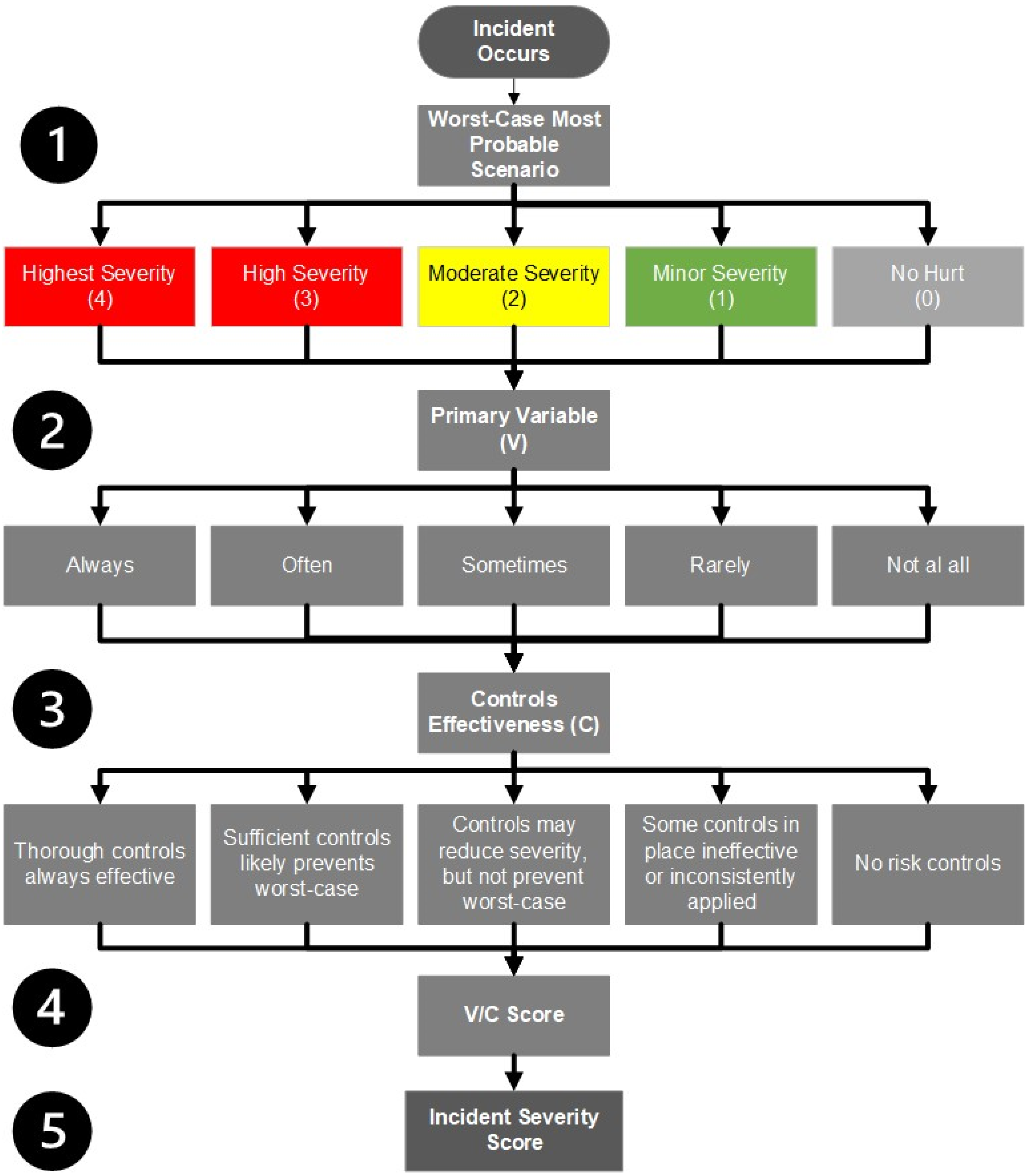

Figure 33. Modified Job Aid for High Severity Incidents

To address this challenge, it requires consistency in decision making, organizational priorities, and resource utilization. A holistic approach to an integrated approach is also a benefit of integrated management systems [54]. However, there are limitations in adopting a graded approach to the response to high-severity events discussed in the conclusion and recommendation section of the thesis.

Additionally, redundancy in incident response can lead to a lack of accountability. To mitigate this, it is recommended to establish cross-functional teams for decision making [54]. From an operational perspective, this may include having an incident severity response team in which various management systems are represented. The team would evaluate incidents to determine potential severity of an incident. While this collaborative approach allows each management system to implement its own mitigation strategies, it still requires top management to approve final decisions. This is especially important in overlap in preventative action determinations.

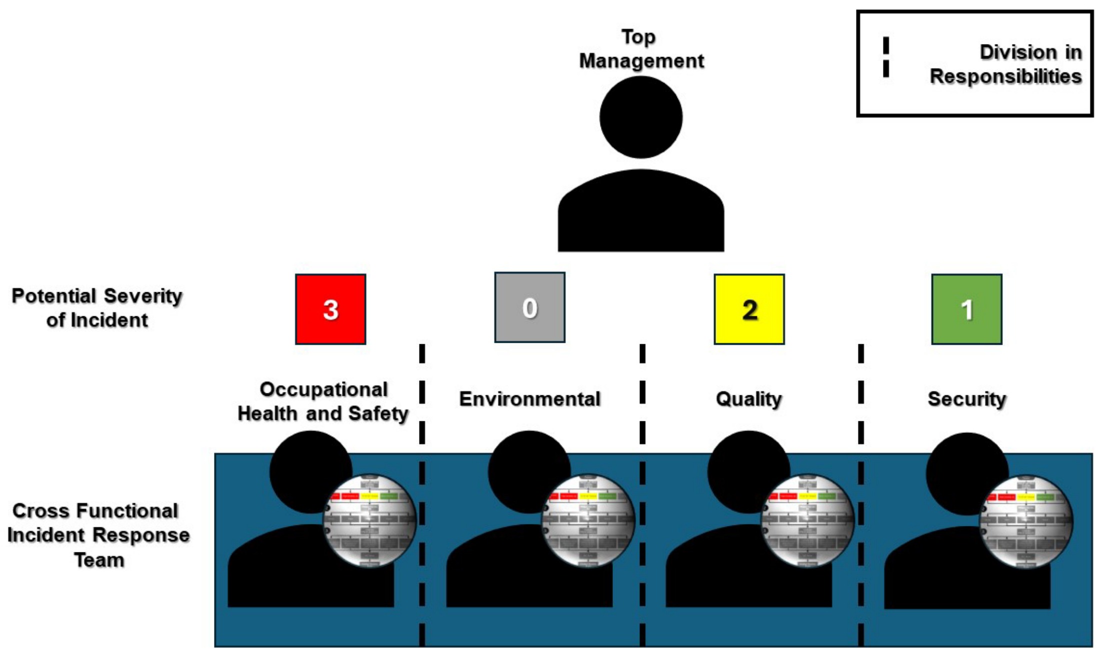

Figure 34. Cross Functional Incident Response Team

A holistic approach to risk management is often necessary due to the overlap in risk to the organization [56]. Siloing of the risk may continue to create fragmented responsibilities. Implementing consistent methodology across management systems reduces the likelihood of unforeseen risks emerging. Figure 18 illustrates individual potential severity impacts within the cross functional team. However, achieving consistency could also be a challenge in a newly formed cross-functional incident severity response team. Consistent and effective risk assessment is always a concern and especially when qualitative measures are employed in risk management [56]. These topics are revisited in the conclusion, although overarching themes expand beyond the theme of this thesis.

Table 20. Example of Potential Severity by Management Systems

Incident Severity Level	Definition	Occupational Health and Safety Examples	Environmental Examples	Quality Examples	Security Examples
Highest Severity (4)	Permanent reputational damage. Significant disruption to operations.	- Fatality or Multiple Fatalities	- Large scale dispersal of hazardous materials	- Inability to complete mission critical procedures	- Theft of nuclear materials
High Severity (3)	Significant damage to reputation. Major disruption to operations.	- Amputation - Loss / impairment of organ functions - Severe to complete loss of hearing	- Release of hazardous gas above regulatory limits	- Critical process failure affecting multiple systems	- Breach of sensitive information
Moderate Severity (2)	Noticeable damage to reputation with impacted operational capabilities available.	- Fractures, loss of tooth/teeth - Partial / single digit amputations - Moderate hearing loss	- Temporary increase in emission within regulatory limits	- Procurement of low-quality materials	- Intentional breach of security protocols
Minor Severity (1)	Minimal reputational damage with little to no disruption to operational capabilities.	- Minor lacerations that bleed freely - Confirmed slight to mild hearing loss - Mild corneal abrasion	- Minor spill of non-hazardous material	- Process deviation that lowers quality of measurements	- Misplaced ID badge without unauthorized access attempt
No Severity (0)	No risk to reputation or disrupt operations.	N/A	N/A	N/A	

LIST OF REFERENCES

[1] U.S. Department of Energy. “About Us.” Department of Energy. Accessed March 31, 2024. https://www.energy.gov/about-us.

[2] U.S. Department of Energy. “National Nuclear Security Administration.” Department of Energy. Accessed March 31, 2024.

[3] Shipman, T. L. 1961. “A Radiation Fatality Resulting from Massive Over-Exposure to Neutrons and Gamma Rays.” In Diagnosis and Treatment of Acute Radiation Injury. Geneva: WHO.

[4] “Serious Injury and Fatality Prevention: Perspectives and Practices.” Campbell Institute. https://www.thecampbellinstitute.org/wp-content/uploads/2018/11/9000013466_CI_Serious-Injury-and-Fatality-Prevention_WP_FNL_single_UPDATED-11-14-18_optimized.pdf (accessed 4 September 2022.

[5] *LLNL at a Glance FACT SHEET*. May 2021 Lawrence Livermore National Laboratory. https://www.llnl.gov/sites/www/files/2021-05/LLNL%20at%20a%20Glance%20FACT%20SHEET.pdf.

[6] Kang, Ho Kyung. 2023. “CORRELATION ANALYSIS OF NAVY FLIGHT MISHAPS.” Thesis, Monterey, CA; Naval Postgraduate School. https://hdl.handle.net/10945/72201.

[7] “Out Front on Airline Safety: Two Decades of Continuous Evolution.” 2018. Newsroom. Federal Aviation Administration. Last Modified August 2, 20218. Accessed March 31, 2024. https://www.faa.gov/newsroom/out-front-airline-safety-two-decades-continuous-evolution#:~:text=The%20commercial%20aviation%20system%20in,fatalities%20per%20100%20million%20passengers.

[8] Heinrich, H. W. 1931. *Industrial accident prevention; a scientific approach.* First edition. New York; London: McGraw-Hill book company, inc., 1931.

[9] Moore, Susan M., Patrick L. Yorio, Emily J. Haas, Jennifer L. Bell, and Lee A. Greenawald. 2020. “Heinrich Revisited: A New Data-Driven Examination of the Safety Pyramid.” *Mining, Metallurgy & Exploration* 37 (6): 1857–1863. https://doi.org/10.1007/s42461-020-00263-0. https://doi.org/10.1007/s42461-020-00263-0.

[10] Šimić, Zdenko, Miguel Peinador Veira, and Reni Banov. 2022. “Correlation between events with different safety significance in nuclear power plants.” Nuclear Engineering and Technology 54 (7): 2510–2518. https://doi.org/ https://doi.org/10.1016/j.net.2022.01.034.

[11] Brown, Jeff. 2020. “Nearly 50 Years of Occupational Safety and Health Data.” U.S. Bureau of Labor Statistics, Last Modified July 2020, accessed March 31, 2024. https://www.bls.gov/opub/btn/volume-9/nearly-50-years-of-occupational-safety-and-health-data.html.

[12] Ezenwa, Sopuruchukwu, Abhijit D. Talpade, Pushkar Ghanekar, Ravi Joshi, Jayachandran Devaraj, Fabio H. Ribeiro, and Ray Mentzer. 2022. “Toward Improved Safety Culture in Academic and Industrial Chemical Laboratories: An Assessment and Recommendation of Best Practices.” *ACS Chemical Health & Safety 29* (2): 202–213. https://doi.org/10.1021/acs.chas.1c00064.

[13] Smith, R. M., and M. L. Jones. 2013. “A Hurt-Based Approach to Safety.” SPE Americas E&P Health, Safety, Security and Environmental Conference.

[14] Jackson, Jay. 2017. “Mining the Diamond: Next Step in the Safety Journey.” AAP/SEG International Conference and Exhibition, London, 15–18 October 2017.

[15] Lara, Raul B., Juliana H. Halbach, Steve Nakasaki, and Sam Y. Paik. “A Case Study in Assessing a Potential Severity Framework for Incidents from a Decadal Sample.” *ACS Chemical Health \& Safety* 0 (0): null. https://doi.org/10.1021/ acs.chas.4c00026.

[16] Hallowell, Matthew, 2023, “Safety Classification and Learning (SCL) Model,” https://www.eei.org/-/media/Project/EEI/Documents/Issues-and-Policy/Power-to-Prevent-SIF/eeiSCLmodel.pdf?la=en&hash= 4E03097C0292F52CB4FA186D0D8CE11876032836.

[17] *Determining Serious Injury and Fatality Potential* (DEKRA, Publication Date not available), https://dekraprod-media.e-spirit.cloud/d6df64f5-b4d9-468c-8ffa-874a167a15ab/media/dekra-or-wp-determining-serious-injury-and-fatality-potential.pdf (accessed March 31, 2024).

[18] Raphael, Moura, Beer Michael, Patelli Edoardo, Lewis John, and Knoll Franz. 2016. “Learning from major accidents to improve system design.” *Safety Science* 84:37-45. doi: https://doi.org/10.1016/j.ssci.2015.11.022.

[19] Remmel, Ariana. 2022. “How to Capture and Use Near-Miss Lab-Incident Reports in Academia.” *ACS Chemical Health & Safety* 29 (2):114-116. doi: 10.1021/acs.chas.2c00016.

[20] “About Us.” U.S. Department of Energy. https://www.energy.gov/ig/about-us.

[21] "Environmental Stewardship and Occupational Health and Safety." Lawrence Livermore National Laboratory (LLNL), accessed March 3, 2024. https://enviroinfo.llnl.gov/.

[22] (ISO), Institutional Organization for Standardization. 2015. ISO 14001:2015 – Environmental Management Systems – Requirements with Guidance for use.

[23] (ISO), International Organization for Standardization. 2018. ISO 45001:2018 – Occupational Health and Safety Management Systems – Requirements with Guidance for Use.

[24] Energy, U.S. Department of. "Occurrence Reporting and Processing System." https://www.energy.gov/ehss/occurrence-reporting-and-processing-system.

[25] 2020. Injury and Illness Reporting Guide. edited by U.S. Department of Energy.

[26] Broussard, C. 2012. DOE O 2314.1B Admin Chg 1, Environment, Safety and Health Reporting. edited by U.S. Department of Energy.

[27] Cournoyer, Michael E., Vincent E. Garcia, Ubaldo F. Gallegos, and Dianne W. Wilburn. 2011. "Investigation of injury/illness data at a nuclear facility." *Journal of Chemical Health & Safety* 18 (5):17-25. doi: 10.1016/j.jchas.2010.11.008.

[28] 2022. "Operating Experience Summary." Department of Energy, accessed 3 January 2024. https://www.energy.gov/sites/default/files/2022-12/OES%202022-05%20CAIRS%20Fiscal%20Year%202022%20In%20Review.pdf.

[29] Cochran, William G. 1977. *Sampling Techniques*. 3rd Edition ed: John Wiley.

[30] Ezenwa, Sopuruchukwu, Abhijit D. Talpade, Pushkar Ghanekar, Ravi Joshi, Jayachandran Devaraj, Fabio H. Ribeiro, and Ray Mentzer. 2022. "Toward Improved Safety Culture in Academic and Industrial Chemical Laboratories: An Assessment and Recommendation of Best Practices." *ACS Chemical Health & Safety* 29 (2):202-213. doi: 10.1021/acs.chas.1c00064.

[31] Dillard, James. 2011. DOE G 450.4-1C, Integrated Safety Management System Guide. edited by Department of Energy.

[32] David, A. Costa, E. Cournoyer Michael, F. Merhege James, E. Garcia Vincent, and N. Sandoval Arnold. 2017. "A primer on criticality safety." *Journal of Chemical Health and Safety* 24 (3): 7–13. https://doi.org/https://doi.org/10.1016/j.jchas.2016.12.001. https://www.sciencedirect.com/science/article/pii/S1871553216301098.

[33] Reason, James. 1997. *Managing the Risks of Organizational Accidents*. 1st Edition ed. London: Routledge. Engineering & Technology.

[34] Larouzée, J. 2017. “Resilience: A New Paradigm of Nuclear Safety: From Accident Mitigation to Resilient Society Facing Extreme Situations.” 257–267. Cham: Springer International Publishing.

[35] DOE-HDBK-1211-2014, Activity-Level Work Planning and Control Implementation. 2014. edited by Department of Energy.

[36] Duane, Kritzinger. 2017. “5 – Failure Modes and Effects Analysis.” In *Aircraft System Safety*, edited by Kritzinger Duane, 101–132. Woodhead Publishing.

[37] Baker, Susan P., and William Haddon. 1974. “Reducing Injuries and Their Results: The Scientific Approach.” *The Milbank Memorial Fund Quarterly. Health and Society* 52 (4): 377–389. https://doi.org/10.2307/3349509.

[38] “Hierarchy of Controls.” Centers for Disease Control and Prevention (CDC), National Institute for Occupational Safety and Health (NIOSH). Accessed March 11, 2024. https://www.cdc.gov/niosh/topics/hierarchy/default.html.

[39] “Identifying Hazard Control Options: The Hierarchy of Controls.” 2023. Occupational Safety and Health Administration (OSHA). Accessed March 11, 2024.

[40] Jeskie, K. B., P. Ashbrook, D. Casadonte, D. Decker, L. J. Doemeny, H. Robert, J. Hill, T. Houts, K. Kretchman, R. Izzo, and et al. 2015. Identifying and evaluating hazards in research laboratories.

[41] Kerns, E., E. A. Masterson, C. L. Themann, and G. M. Calvert. 2018. “Cardiovascular conditions, hearing difficulty, and occupational noise exposure within U.S. industries and occupations.” *Am J Ind Med* 61 (6): 477–491. https://doi.org/10.1002/ajim.22833.

[42] Medicine, Institute of. 2006. *Noise and Military Service: Implications for Hearing Loss and Tinnitus*. Edited by Larry E. Humes, Lois M. Joellenbeck and Jane S. Durch. Washington, DC: The National Academies Press.

[43] National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Identifying Disabling Medical Conditions Likely to Improve with Treatment.” In Musculoskeletal Disorders. Washington (DC): National Academies Press (US). Available from: https://www.ncbi.nlm.nih.gov/books/NBK559512.

[44] SangUk, Han, Saba Farzaneh, Lee SangHyun, Mohamed Yasser, and Peña-Mora Feniosky. 2014. “Toward an understanding of the impact of production pressure on safety performance in construction operations.” *Accident Analysis & Prevention* 68: 106–116. https://doi.org/https://doi.org/10.1016/j.aap.2013.10.007. https://www.sciencedirect.com/science/article/pii/S0001457513004041.

[45] Harrell, Jr Frank E. 2015. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd 2015. ed. *Springer Series in Statistics*. Cham: Springer International Publishing.

[46] Agresti, Alan. 2002. *Categorical data analysis*. 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, New Jersey: Wiley-Interscience.

[47] Seabold, Skipper, and Josef Perktold. 2010. "statsmodels: Econometric and Statistical Modeling with Python." Proceedings of the 9th Python in Science Conference. https://www.statsmodels.org.

[48] McKinney, Wes. 2010. "pandas: Powerful Python Data Analysis Toolkit." https://pandas.pydata.org/.

[49] Fletcher, R. 1988. *Practical Methods of Optimization*. 2nd ed. Newark: John Wiley & Sons, Incorporated.

[50] "DOE OPEXShare." 2024. https://doeopexshare.doe.gov/.

[51] Energy, U. S. Department of. 2024. "National Environmental Policy Act." https://www.energy.gov/em-la/national-environmental-policy-act.

[52] (ISO), International Organization for Standardization. 2015. ISO 9001:2015 – Quality Management Systems – Requirements.

[53] Juran, J. M. ed, and Frank M. Gryna. 1988. *Juran's quality control handbook*. 4th ed. San Francisco, CA: McGraw-Hill.

[54] (ISO), International Organization for Standardization. 2018. ISO Handbook 2018–11 – The Integrated Use of Management System Standards (IUMSS).

[55] Sui, Yang, Rui Ding, and Hanqing Wang. 2018. "An integrated management system for occupational health and safety and environment in an operating nuclear power plant in East China and its management information system." *Journal of Cleaner Production* 183: 261–271. https://doi.org/https://doi.org/10.1016/j.jclepro.2018.02.101. https://www.sciencedirect.com/science/article/pii/S095965261830413X.

[56] Hubbard, Douglas W. 2020. *The failure of risk management: why it's broken and how to fix it*. Second edition. ed. Hoboken, New Jersey: Wiley.

[57] Hunter, J.D. Matplotlib: Visualization with Python. 2007. https://matplotlib.org/.

[58] Harris, C.R. et al. NumPy: The fundamental package for scientific computing with Python. 2020. https://numpy.org/.

[59] Seaborn: Statistical data visualization. 2021. https://seaborn.pydata.org/.

[60] Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. 2011. https://scikit-learn.org/.

[61] Project Jupyter: Open tools for interactive computing. 2016. https://jupyter.org/.

[62] 2023. "NIST/SEMATECH e-Handbook of Statistical Methods: 1.3.6.6.19. Poisson Distribution." https://www.itl.nist.gov/div898/handbook/eda/section3/eda366j.htm.

[63] Blood, C. G., and E. R. O'Donnell. 1995. "A system to project injury and illness incidence during military operations." J Med Syst 19 (6): 457–64. https://doi.org/10.1007/bf02260849.

[65] Nicholson, Alan, and Yiik-Diew Wong. 1993. "Are accidents poisson distributed? A statistical test." Accident Analysis & Prevention 25 (1): 91–97. https://doi.org/https://doi.org/10.1016/0001-4575(93)90100-B. https://www.sciencedirect.com/science/article/pii/000145759390100B.

[66] Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. "Model Assessment and Selection." In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, edited by Trevor Hastie, Robert Tibshirani and Jerome Friedman, 219–259. New York, NY: Springer New York.

[67] Akaike, H. 1974. "A new look at the statistical model identification." *IEEE Transactions on Automatic Control* 19 (6): 716–723. https://doi.org/10.1109/TAC.1974.1100705.

[68] Gideon, Schwarz. 1978. "Estimating the Dimension of a Model." *The Annals of Statistics* 6 (2): 461–464. https://doi.org/10.1214/aos/1176344136. https://doi.org/10.1214/aos/1176344136.

[69] SciPy Developers. 2023. "scipy.special.expit." https://docs.scipy.org/doc/scipy/reference/generated/scipy.special.expit.html.

[70] *Industrial Safety Guidelines for Nuclear Facilities*. 2018. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY.

[71] *Risk Aggregation for Nuclear Installations*. 2021. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY.

[72] Li, Z., M. Yao, Z. Luo, X. Wang, T. Liu, Q. Huang, and C. Su. 2024. "A chemical accident cause text mining method based on improved accident triangle." *BMC Public Health* 24 (1): 39. https://doi.org/10.1186/s12889-023-17510-w.

[73] Campesato, Oswald. 2023. "Metrics in Statistics." 75–104. United States: Mercury Learning & Information.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Fort Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California