



Construction Industry Institute®

Precursor Analysis for the Construction Industry: A Systematic Method for Predicting and Preventing Fatal and Disabling Injuries

Matthew Hallowell

Dillon Alexander

University of Colorado at Boulder

John Gambatese

Oregon State University

Construction Industry Institute

Abbott	AECOM
Ameren Corporation	Affiliated Construction Services Inc.
American Transmission Company LLC	AMEC Foster Wheeler
Anadarko Petroleum Corporation	Atlas RFID Solutions
Anheuser-Busch InBev	Autodesk, Inc.
Aramco Services Company	AZCO INC.
ArcelorMittal	Baker Concrete Construction Inc.
Architect of the Capitol	Barton Malow Company
Ascend Performance Materials	Bechtel Group, Inc.
AstraZeneca	Benham Constructors, LLC
BP America, Inc.	Bentley Systems Inc.
Cargill, Inc.	Bilfinger Industrial Services Inc.
Chevron	Black & Veatch
ConocoPhillips	Burns & McDonnell
Consolidated Edison Company of New York	CB&I
The Dow Chemical Company	CCC Group
DTE Energy	CDI Corporation
Eastman Chemical Company	CH2M
Eli Lilly and Company	Construtora Norberto Odebrecht S.A.
Enbridge Inc.	Coreworx Inc.
EnLink Midstream	CSA Central, Inc.
ExxonMobil Corporation	Day & Zimmermann
General Electric Company	Eichleay, Inc.
General Motors Company	Emerson
GlaxoSmithKline	Ensto, Inc.
Global Infrastructure Partners	eProject Management, LLC
Greater Toronto Airports Authority	Faithful+Gould
Honeywell International Inc.	Fluor Corporation
Huntsman Corporation	Hargrove Engineers + Constructors
Irving Oil Limited	Hilti Corporation
Kaiser Permanente	I.M.P.A.C.T.
Koch Industries, Inc.	IHI E&C International Corporation
LyondellBasell	IHS
Marathon Petroleum Corporation	Jacobs
National Aeronautics & Space Administration	JMJ Associates LLC
NOVA Chemicals Corporation	JV Driver Projects Inc.
Occidental Petroleum Corporation	KBR
ONEOK, Inc.	Kiewit Corporation
Ontario Power Generation	Lauren Engineers & Constructors, Inc.
Pacific Gas and Electric Company	Matrix Service Company
Petroleo Brasileiro S/A - Petrobras	McCarthy Building Companies, Inc.
Petronas	McKinsey & Company, Inc.
Phillips 66	Midwest Steel, Inc.
Pioneer Natural Resources	Parsons
Potash Corporation of Saskatchewan	Pathfinder, LLC
The Procter & Gamble Company	PCL Constructors, Inc.
Public Service Electric & Gas Company	PricewaterhouseCoopers Advisory Services LLC
Reliance Industries Limited (RIL)	PTAG, Inc.
SABIC - Saudi Basic Industries Corporation	Quality Execution, Inc.
Shell Global Solutions US Inc.	Richard Industrial Group
Smithsonian Institution	The Robins & Morton Group
Southern Company	S & B Engineers and Constructors, Ltd.
Tennessee Valley Authority	SBM Offshore
Tesoro Corporation	Skanska USA
TransCanada Corporation	SNC-Lavalin Inc.
U.S. Army Corps of Engineers	Supreme Group
U.S. Department of Commerce/NIST/EL	Technip
U.S. Department of Defense/ Tricare Management Activity	UniversalPegasus International
U.S. Department of Energy	Victaulic
U.S. Department of State	Wanzek Construction, Inc.
U.S. Department of Veterans Affairs	The Weitz Company, Inc.
U.S. General Services Administration	Wilhelm Construction, Inc.
The Williams Companies, Inc.	Wood Group Mustang
	WorleyParsons
	Yates Construction
	Zachry Group
	Zurich



PRECURSOR ANALYSIS FOR THE CONSTRUCTION INDUSTRY: A SYSTEMATIC
METHOD FOR PREDICTING AND PREVENTING FATAL AND DISABLING INJURIES

by

Matthew Hallowell
Dillon Alexander
University of Colorado at Boulder

John Gambatese
Oregon State University

A Report to
Construction Industry Institute
The University of Texas at Austin

Under the Guidance of
CII Research Team 321
Using Precursor Analysis to Prevent Low-frequency High-impact Events,
Including Fatalities

CII Research Report 321-11
October 2016

© 2016 Construction Industry Institute®

The University of Texas at Austin.

CII members may reproduce and distribute this work internally in any medium at no cost to internal recipients. CII members are permitted to revise and adapt this work for their internal use, provided an informational copy is furnished to CII.

Available to non-members by purchase; however, no copies may be made or distributed and no modifications may be made without prior written permission from CII. Contact CII at <http://construction-institute.org/catalog.htm> to purchase copies. Volume discounts may be available.

All CII members, current students, and faculty at a college or university are eligible to purchase CII products at member prices. Faculty and students at a college or university may reproduce and distribute this work without modification for educational use.

Printed in the United States of America.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
2. RESEARCH OBJECTIVES	3
3. BACKGROUND	4
Phase 1 Background: Assessing the extent to which energy magnitude predicts injury.....	4
severity	
<i>Energy release theory.....</i>	4
<i>Comparison to extant methods of injury propagation</i>	4
<i>Parallels between energy release theory and natural disaster research</i>	5
<i>Proposed new theory: Energy-based safety risk analysis</i>	5
Phase 2 Background: Identifying precursors and creating a precursor analysis protocol.....	7
<i>Precursor analysis programs in other industries</i>	8
<i>Precursor analysis in construction</i>	9
Phase 3 Background: Testing the protocol through an experiment, and independent	10
validation	
Phase 4 Background: Creating and validating an objective predictive model	10
<i>Safety risk analysis.....</i>	10
<i>Safety leading indicators.....</i>	11
<i>Predictive analytics for safety</i>	12
4. HYPOTHESES AND RESEARCH QUESTIONS	14
Phase 1 Hypotheses: Assessing the extent to which energy magnitude predicts injury	14
severity	
Phase 2 Questions: Identifying precursors and creating a precursor analysis protocol	14
Phase 3 Hypothesis: Testing the protocol through an experiment, and independent	14
validation	
Phase 4 Hypothesis: Creating and validating an objective predictive model.....	14
5. RESEARCH METHODS	15
Phase 1 Methods: Assessing the extent to which energy magnitude predicts injury severity ...	15
Phase 2 Methods: Identifying precursors and creating a precursor analysis protocol	17
<i>Creating a complete catalog of potential precursors and subsequent data collection protocol</i>	18
Phase 3 Methods: Testing the protocol through an experiment and independent validation ...	19
<i>Step 1: Gather and screen cases</i>	20
<i>Step 2: Expert team assesses cases and make predictions</i>	21
<i>Step 3: Validate with diverse groups</i>	21
Phase 4 Methods: Creating and validating an objective predictive model.....	22
<i>Step 1: Gather structured data via interpretations of the precursor analysis protocol.....</i>	22
<i>Step 2: Eliminate redundancy in initial data collection instrument</i>	23
<i>Step 3: Use principal components analysis to objectively reduce the dimensions of the dataset</i>	25

<i>Step 4: Build predictive model using principal components and generalized linear modeling</i>	25
<i>Step 5: Validate predictive model using results from independent case studies</i>	26
6. RESULTS	33
Phase 1 Results: Assessing the extent to which energy magnitude predicts injury severity	33
<i>Testing statistical significance among all injury severity levels</i>	33
<i>Defining a “high energy” threshold</i>	35
<i>Explaining the variability and extreme values</i>	36
<i>Defining ‘high energy’ hazards</i>	37
Phase 2 Results: Identifying precursors and creating a precursor analysis protocol	39
Phase 3 Results: Testing the protocol through an experiment and independent validation	44
Phase 4 Results: Creating and validating an objective predictive model	46
<i>Principal components analysis</i>	46
<i>Generalized linear model</i>	47
<i>Predictive skill of the generalized linear model</i>	47
7. CONCLUSIONS AND RECOMMENDATIONS	50
Phase 1 Conclusions: Assessing the extent to which energy magnitude predicts injury severity	50
Phase 2 Conclusions: Identifying precursors and creating a precursor analysis protocol	51
Phase 3 Conclusions: Testing the protocol through an experiment, and independent validation	52
Phase 4 Conclusions: Creating and validating an objective predictive model	52
8. KEY FINDINGS	53
APPENDIX A – EXAMPLE CASE	55
APPENDIX B – SCORECARD FOR EXAMPLE CASE	58
REFERENCES	59

1. INTRODUCTION

In 2015 the United States Bureau of Labor Statistics (BLS) released the 2013 census data for work-related fatalities, injuries, and illnesses (BLS 2015). Construction recorded the highest number of fatalities of any industry, with a total of 856. Tragically, this marks the highest number of fatalities in the construction industry since 2009 and, more importantly, highlights the growing concern that the fatality rate in construction has recently plateaued. As shown in Figure 1, similar observations have been made of Construction Industry Institute (CII) data. As these data suggest, there is a need for new injury prevention methods that focus purely on the fatalities that continue to plague the industry. Enhancing traditional safety methods and using leading information to predict and prevent catastrophic fatalities is a logical next step for the industry.

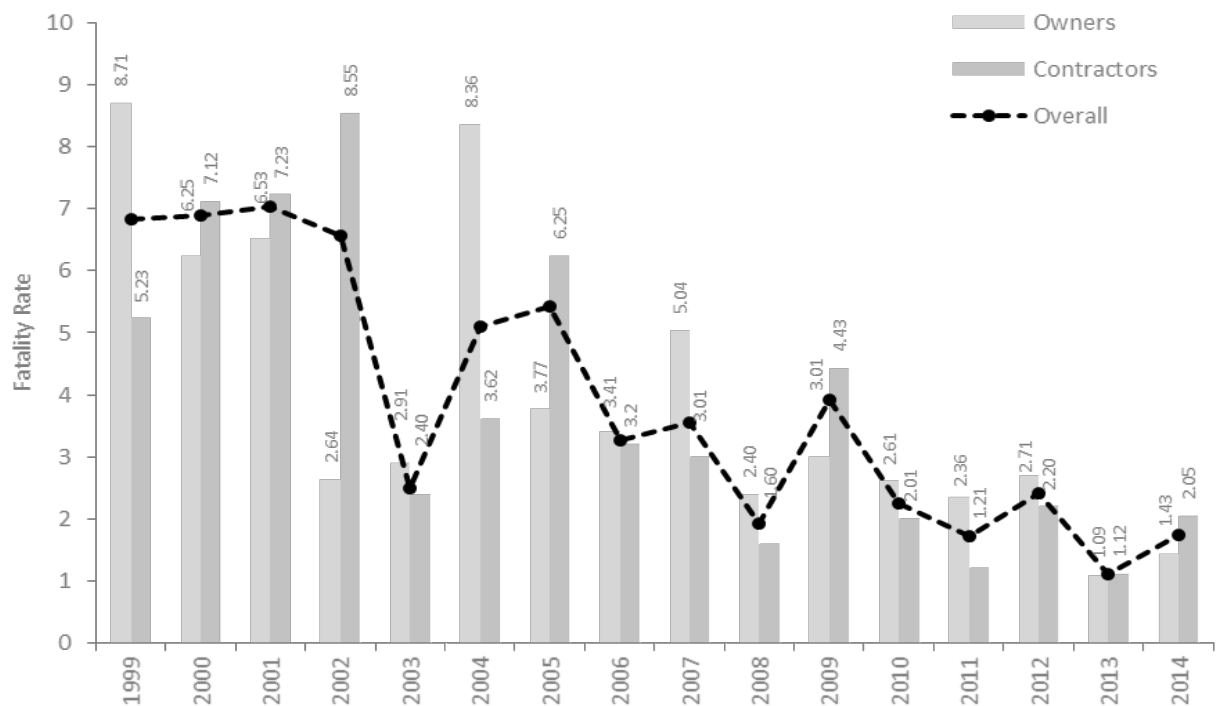


Figure 1 – CII-reported fatality statistics

One of the challenges of analyzing fatal and disabling injuries is that they (fortunately) occur infrequently in comparison to less severe events. The infrequent nature of fatal events makes it especially difficult to identify latent factors and analyze causal trends. Researchers have expressed frustration in this regard, questioning why devastating fatal incidents continue to occur in the construction industry given the amount of investigation and analysis that has taken place and our current level of safety knowledge (Wu et al. 2010). Fortunately, other industries that rely on the prevention of infrequent and catastrophic events have found a successful strategy to manage rare, high-impact events: precursor analysis.

Similar to the National Academy of Engineering (2004), we define a precursor as a *reasonably detectable event, condition, or action that serves as warning signs to a fatal or disabling injury*. Precursor analysis was first introduced following the Three Mile Island nuclear plant meltdown of 1979 (Minarick, 1990). Other industries such as aviation, chemical manufacturing, and pharmaceuticals, and the National Aeronautical and Space Agency's (NASA) space program have all since created specific precursor analysis programs (National Academy of Engineering, 2004). Although precursors from some industries may not be directly translatable to other industries, the precursor identification and analysis processes are

fundamentally similar. Further, since we are focused on high-impact, low-frequency (HILF) events, we define them as *high-energy events that result in or have the potential to result in a fatality or life-altering injury or illnesses and precursors as reasonably detectable events, conditions, or actions that serve as warning signs of an event.*

Our goal in the present study was to create an empirically validated precursor analysis method for HILF events. In pursuit of this goal, we established the following objectives (See Figure 2):

1. Identify and document causal factors of construction fatalities;
2. Translate the list of factors into a series of leading questions that can be asked in the field;
3. Apply the precursor questionnaire to collect responses for cases of successful work (no incident) and failed work (fatality, injury, or near miss);
4. Conduct a controlled experiment to measure the extent to which the data collected via the protocol and judgment can be used together in a leading and predictive fashion to distinguish between success and failure; and
5. Measure the generalizability of the method by repeating the experiment with multiple groups with varying degrees of industry experience.
6. Create a mathematical model that objectively predicts the probability that a work scenario will result in a HILF event.
7. Validate the predictive model using new cases.

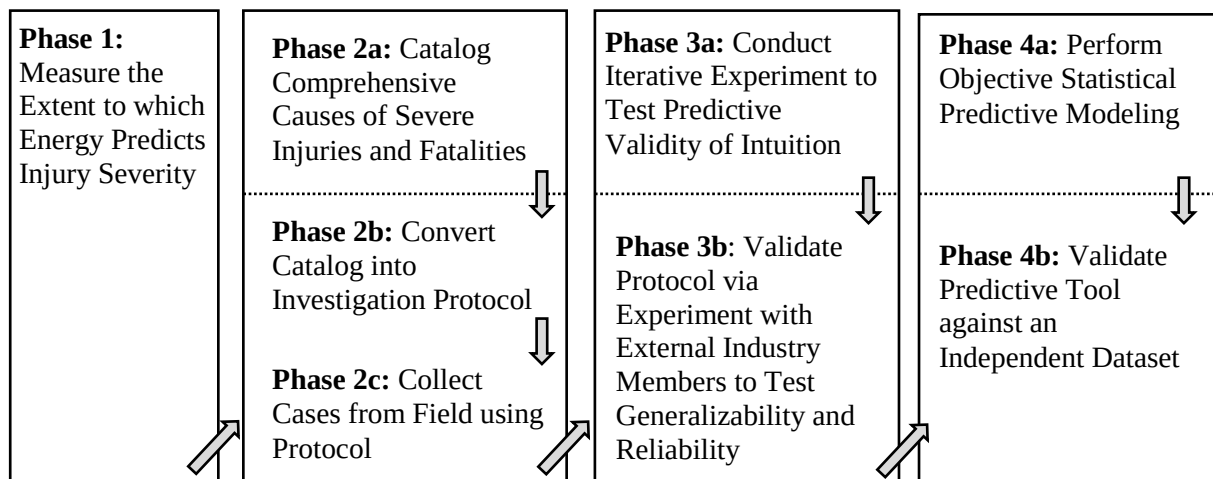


Figure 2 – Overarching research process and organization of this summary

2. RESEARCH OBJECTIVES

This research is a step toward the long-term goal of creating a precursor analysis method that can be reasonably applied in a leading fashion to identify anomalous work situations that have extraordinary potential for a fatal event.

We build upon but deviate from past methods of proactive safety analysis. First, we take a new approach to the characterization and collection of the independent predictor variables. Specifically, the leading data were collected via a brief but targeted conversation with workers *prior* to a work period rather than observations or accounts collected *after* an incident. Second, we make the first attempt to objectively distinguish success (i.e., completion of work with no unwanted release of energy) from failure (i.e., work with an unwanted release of energy that either did or had the potential to cause significant injury or death). Because all past safety prediction models have been built using reports of injuries, they were all limited to predicting the severity or type of outcome *if one were to occur* (e.g., Esmaeili et al. 2015b; Tixier et al. 2016). This work deviates by including cases of both success and failure and creating models that forecast the probability that a fatality or disabling injury *will occur at all*.

3. BACKGROUND

This research study was extensive and multi-faceted. Therefore, the background necessary to properly frame the work is extensive and diverse. Here, we provide a detailed review of literature, highlighting key theory associated with each phase of the research. When appropriate, we combine phases for efficiency.

Phase 1 Background: Assessing the extent to which energy magnitude predicts injury severity

Applying the concept of energy to explain variability in the severity of construction injuries requires a description of our epistemological positions on both safety risk analysis and safety energy. Also, we must explain the scientific basis for our proposed extension of the energy theory. Thus, we dedicate this background to reviewing perspectives of and approaches to safety risk analysis, the theory of safety energy, extant methods of explaining accident propagation, and comparable scientific approaches to risk modeling from the field of natural disaster research. We use this background to frame our theoretical point of departure and our contributions to the body of knowledge.

Energy release theory

The inspiration for an energy-based approach to safety risk analysis originated with Haddon's (1968) etiologic basis for describing injury severity. Using medical diseases and symptoms as an analogy, Haddon (1968) argued that injuries are not caused by the seemingly infinite and random collaboration of causal factors that describe the specific circumstances of the accident (i.e., symptoms). Rather, he claims that injuries should be defined by their fundamental cause, hazard energy (i.e., disease), or more specifically, the release of hazard energy and contact by an individual. Under this etiologic explanation, there must be a form of energy exchange in excess of the body's vulnerability in order for an injury to be sustained. Approximately ten years later, Haddon (1980) further expanded upon this etiological theory by introducing the concept that injury prevention should focus on the removal, reduction, isolation, or control of hazard energy.

Fleming (2009) contributed to the theory by defining various energy sources that cause injury. His argument was that industry application of the energy theory requires focus on the management of specific types of hazard energy sources (e.g., gravity, motion, electrical). Most recently, principles of the energy release theory were organized into a ten-energy-source mnemonic, which was then field-tested in an effort to measure its impact on hazard recognition skills (Albert et al. 2014a;b;c). The results of the multiple baseline test indicate that the mnemonic based learning increases hazard recognition skills by approximately 30% when compared with traditional safety planning activities alone (e.g., job safety analyses and checklists). Despite these significant advancements, there is still a dearth of research that investigates the scientific extension and practical application of hazard energy within occupational safety.

Comparison to extant methods of injury propagation

The energy release theory is different from more traditional methods of explaining accident propagation. There is a relatively large body of knowledge related to accident causation, which aims to explain the reasons *why* injuries occur. In the form of risk analysis, some researchers have begun to adopt a position that the risk of injury is related to the physical characteristics of the work environment and tasks (e.g., Esmaeili et al. 2015; Tixier et al. 2016). Others have focused on the role that specific entities play in accident causation like designers (e.g. Gambatese et al. 2008), clients (e.g., Huang and Hinze 2006), supervisors (e.g., Hardison et al. 2015), and workers (e.g., Rasmussen et al. 1990). In a more inclusive epistemological position, accident causation researchers agree that injuries are the result of the confluence of system failures (Suraji et al 2001; Haslam et al. 2005). This position emanates from theories like the *Swiss Cheese Model* (Reason 1990). This entire body of knowledge focuses on what causes an injury to occur rather than how severe the injury might be if one were to occur. Here we focus exclusively on predicting and explaining injury severity. This defines our theoretical point of departure.

It should be noted that there have been a few studies that aimed to predict injury severity. For example, Esmaeili et al. (2015) attempted to predict injury severity using the presence or absence of fundamental attributes of the work environment and a generalized linear model. Their position was that injury severity could be predicted by the tools, equipment, tasks, environmental conditions, and other physical attributes of the workspace. The results indicated that the model was able to predict the type of injury (e.g., fall or struck by) but was not significantly better than random at predicting injury severity. Tixier et al. (2016) confirmed these results using similar data and machine learning techniques. We offer an alternative approach to severity prediction using a scientifically fundamental position that the quantity and magnitude of energy is the true predictor. This position was inspired by the fields of disaster research as discussed below.

Parallels between energy release theory and natural disaster research

Our concept of energy-based safety risk analysis parallels modern theory of natural hazard risk assessment techniques. In essence, predicting the impact of a natural hazard relies on three fundamental factors: 1) the potential for a natural hazard; 2) the natural hazard's impact intensity; 3) and the vulnerability of the affected community (Lindell & Prater, 2003). Similar to occupational injuries, the impact (severity) of a natural hazard is very difficult to predict because it may be comprised of several hazards that may inflict damage on a particular area (Lindell & Prater, 2003). For example, one agent, such as a hurricane, may inflict casualties and damage through wind, rain, storm surge, and inland flooding (Bryant, 1991). For this reason, the impact intensity of a natural hazard can generally be defined in terms of the physical materials involved (i.e. liquid in a flood or solid material in a landslide) and the energy these materials impart (Lindell & Prater, 2003). Whether the goal is to predict the potential impact of landslides using the mass and velocity of a potential landslide area (Cardinali et al., 2002) or predict the potential impact of a hurricane by measuring wind speed (Moon, et al., 2007), using a natural hazard's energy allows the potential impact of various natural hazards to be expressed in one universally relatable metric (Lindell & Prater, 2003). We postulate that the same theory holds true for safety, where many sources of energy can contribute to the potential severity of an event.

The second portion of any natural hazard risk assessment is the vulnerability of an affected community and the community's ability to withstand a natural hazard. The majority of natural hazard risk analysis research involves the study of vulnerability. For example, several studies have created models to assess the vulnerability of potentially affected areas based on variables ranging from a country's Gross Domestic Product (Kahn, 2005), the money invested in mitigation and emergency practices in the area (Lindell & Prater, 2003), the density of population and complexity of commercial infrastructure (Wei, Fan, Lu, & Tsai, 2004), or the area's reliance on the surrounding ecosystem (Tang, Li, Lei, Wang, & Shen, 2015).

Proposed new theory: Energy-based safety risk analysis

From elemental concepts of energy release theory, we postulate that: *the severity of a potential injury is determined by the ratio of energy intensity and the vulnerability of the human body part to which the energy is transferred.* Furthermore, a hazard's energy intensity is defined as *the ratio of the hazard's energy magnitude and the mechanism by which energy contacts the human body.* The likelihood component of risk is, in turn, defined as the human, managerial, political, social, and stochastic factors that define the probability that energy will be released and will contact a worker. Figure 1 illustrates the proposed relationships among these factors organized in to parallel prevailing natural disaster models presented by Lindell and Prater (2003). Here, the energy transfer mechanics refer to the shape and size of the object possessing energy, which may contact the human body (e.g., a sack of concrete, tape measure, mobile equipment).

Because the vulnerability of the human body is highly variable among subjects and difficult to observe in a leading manner, the focus of this paper will be on hazard energy magnitude and intensity. Specifically, we aim to assess the extent to which energy magnitude and energy intensity predicts injury severity. Since

we postulate that energy intensity is a better predictor of injury severity, we show this relationship in Figure 1.

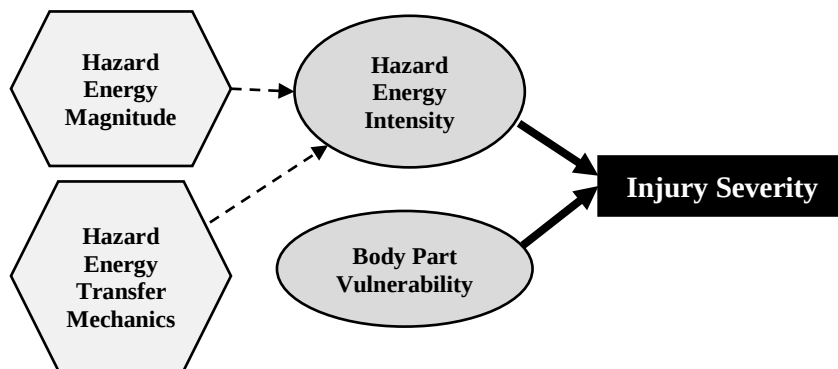


Figure 3: Proposed Energy-Based Occupational Safety Risk Analysis Framework

Hazard energy magnitude

In order to compute energy intensity, one must first measure *energy magnitude*. Since a variety of energy sources exist in occupational environments (e.g., gravitational, kinetic, radiation), there are various methods of computing energy magnitude. For example, using basic physics, gravitational potential energy (PE) can be directly measured by finding the product of the mass of an elevated object, the height of the object, and the gravitational constant (Equation 1).

$$PE = mgh \quad \text{Equation 1}$$

where m = mass; g = gravitational constant; and h = height that the object or subject could fall.

In a practical sense, the computation of PE is quite simple as it only requires the measurement of height and mass along with the knowledge that when either increases, the quantity of PE increases proportionally. However, translating energy magnitude to potential injury severity can be abstract without formal analysis. Take, for example, the conditions where both a 0.5 kg tape measure and a 7 kg sledgehammer are dropped from a height of 2 meters. It is obvious that, due to the difference in mass, the sledgehammer possesses a greater quantity of energy than the tape measure. Also, it may seem intuitive that the sledgehammer has the potential to inflict greater damage and result in a more severe injury. However, it becomes comparatively difficult to intuit energy magnitude and make comparisons in extremes when more than one variable changes at a time or when different energy types are compared. For example, assessing the comparative energy between a 0.5 kg tape measure dropped from 10 meters and a 7 kg sledgehammer dropped from 2 meters is not as easy. In such cases, formal energy computations with consistent units and reliable predictive models could allow for an objective assessment of risk. In other words, formal analyses would help a person to evaluate the quantity of the energy and interpret the risk associated with the energy without relying on judgment that is subject to cognitive biases.

Hazard energy intensity

Synonymous with natural hazard risk analysis, we postulate that the actual danger associated with a hazard energy source depends not only on energy magnitude but also on the physical characteristics of the hazard itself. In order to be theoretically transferable to any hazard energy source, the term ‘relation’ used within the definition of energy intensity references two scenarios: 1) a scenario in which a greater transfer mechanism value will increase the likelihood an individual is severely injured and be expressed in a unit of time (i.e. radiation, sound, and chemical); 2) and a scenario in which a smaller energy transfer mechanism value will result in a more severe injury and be expressed in a unit of area (i.e. pressure, gravity, and motion). Equations 2 and 3 present the respective relationships above scenarios, respectively. From a

scientific standpoint, the second scenario is analogous to pressure and the translation of energy to force through Newton's Second Law of Motion. The second scenario will be the focus of the experimental portion of this study; however, the proposed theory can apply to any hazardous energy source.

*Hazard Intensity = (Energy Magnitude) * (Energy Transfer Period)* Equation 2

$$\text{Hazard Intensity} = \frac{\text{Energy Magnitude}}{\text{Energy Transfer Area}} \quad \text{Equation 3}$$

To better illustrate the relationship between energy intensity and the severity of a sustained injury for the second scenario (Equation 3), a comparison between a 0.5 kg tape measure and a 0.5 kg concrete chisel dropped from 3 meters provides a compelling example. The concrete chisel has a minimum contact area at its tip of approximately 0.6 cm² while a standard tape measure on its smallest edge is approximately 12.9 cm². Assuming that both objects strike a person with the identical impact duration and velocity (i.e., 0.05 seconds and 7.67 m/s, respectively), the pressure of the tape measure and concrete chisel are 59.5 kPa and 1279 kPa, respectively. Although the tape measure and concrete chisel possess an equivalent amount of potential gravitational energy (14.7 joules), the disparity in impact area between the objects causes dramatic differences in pressure and injury.

Because the impact duration needed to calculate the rate of force transfer (i.e. impulse) cannot be readily measured in a leading fashion, we simplify the computation to energy intensity, which, for the example, would be 1.14 joules/cm² for the tape measure and 24.5 joules/cm² for the concrete chisel. The concept of energy intensity provides a reasonable compromise between the scientific validity and practicality, using variables that are readily observed in the field.

Vulnerability of the human body

Returning to the analogy of natural hazard modeling, the magnitude of a natural disaster, such as a hurricane, does not fully predict the magnitude of damage. For example, due to differences in infrastructure, a hurricane striking the coast developing country would presumably inflict more human casualties than a hurricane of the same magnitude striking a developed country. This concept, referred to as vulnerability, also applies in the context of energy-based safety risk where the extent of injury caused by a particular hazard varies depending upon the vulnerability of a particular body part to which the energy is transferred. For example, in the unfortunate case of the New Jersey construction worker being struck by a falling tape measure (Santora, 2014), the injury most likely would have not been fatal had the tape measure struck the worker in the shoulder rather than his head. Additionally, personal protective equipment (PPE) items, such as a hard hat or padded gloves, can improve the body's resistance to injury and reduce vulnerability to impact by dispersing the energy across its surface. Although we theorize that the vulnerability of the impacted body part heavily influences the severity of injury suffered, it is difficult to evaluate this factor in a predictive sense due to the speculative nature of estimating variables such as the variations in worker actions and body position. These values are often impossible to quantify accurately from second-hand accounts in injury reports. For these reasons, vulnerability was omitted from analysis with the caveat that inclusion of this variable in future research may improve the overall scientific validity of the theory.

Phase 2 Background: Identifying precursors and creating a precursor analysis protocol

The purpose of this background section is to describe the precursor analysis process implemented by industries that have developed and maintained successful programs and to describe the state of precursor analysis in construction. We relied heavily on reports published for NASA's Accident Precursor Analysis (APA) program, the aviation industry's Aviation Safety Action Program (ASAP), and the nuclear industry's accident sequence precursors (ASP) program. We also provide a targeted review of academic literature that highlights the causes of construction fatalities (see Table 1).

Precursor analysis programs in other industries

Each industry exhibits a distinctive set of precursors that emanate from the unique tasks, environments, and risks present in the industry. For example, reflecting the complexity of the nuclear energy production and distribution system, the ASP program includes 422 individual precursors that relate to nuclear core damage (Minarick, 1990). Although NASA must also manage a complex system, their precursors are very different and focus on the malfunction of system components (NASA/Sp-2011-3423, 2011). Further, aviation precursors relate more to human error, noncompliance, high tolerance of risk, lack of hazard recognition, inadequate management procedures, and lack of competence (Wiegmann and Shappell, 2001). Despite the differences in precursors among industries, the method of creating each program was surprisingly similar. The creation process can be sequenced into four fundamental steps: (1) conducting an in-depth deterministic analysis of past incidents to identify potential precursors; (2) building a system for incident investigation with focus on suspected precursors; (3) conducting a probabilistic risk assessment (PRA) for continuously updated datasets; and (4) building an information feedback loop for continuous improvement. Although the aviation, transportation, medical, chemical manufacturing, and space exploration industries have all modified the original ASP program to their specific needs, the four fundamental steps for building a precursor analysis program remained consistent.

Component 1: Identify precursors from past events using deterministic event analysis

The creation of a precursor analysis system begins by leveraging historical information from events to identify potential precursors of future events. To achieve this, both NASA and the nuclear industry incorporated deterministic event analysis of past incidents. Deterministic event analysis is the process of identifying causal factors and emerging trends by examining the circumstances of past events (ETSON 2013). To begin, root cause analysis is applied to identify precursors by analyzing the type of human error involved, whether the procedure was adequate or followed correctly, characteristics of the working environment, the qualifications of those involved, and the presence or absence of organizational barriers (IAEA, 2004). Then, Reason's (1990) *Swiss Cheese* model is typically used to organize and layer the precursors into an investigation protocol. Once candidate precursors are identified and prioritized, industries begin to formally collect and analyze a large volume of targeted precursor data to uncover patterns and latent causes.

Component 2: Incident investigation with specific information criteria

Detailed data drive all successful precursor programs. When collecting data required for precursor analysis, successful industries have established comprehensive incident investigation protocols that create homogeneity and ensure standard quality of the data. For instance, the ASP program relies on analyzing reports of undesired events, which include key criteria such as the description of the complete event, the event sequence, a determination of deviations from standard protocol, root cause analysis, actual or potential consequences, and future corrective actions (ETSON, 2013; IAEA, 2004). All nuclear power plants participate in the ASP program and event reports are submitted to the regulatory body for further review in a standardized format (IAEA, 2004). Only events with severe or potentially severe consequences are considered for further analysis (Johnson and Rasmuson, 1996). NASA's APA program also includes an investigation system that uses reports for all anomalous events. These reports must follow established information criteria when submitted to the NASA database and are subsequently evaluated by the APA committee if anomalous events are deemed to have potentially severe consequences (NASA/Sp-2011-3423, 2011). Similarly, the aviation industry has developed and implemented a standardized but voluntary program for the submission of incidents (DOT FAA/AFS-230, 2002; DOT FAA/AFS-800, 2006).

It is important to note that, because catastrophic events are rare, analyzing only high impact cases would not yield a sufficient dataset for any pattern recognition or advanced diagnostic statistics. To address this barrier, operational data from potentially severe incidents (i.e., potentially high-impact "near misses" or "close calls") are typically leveraged to develop a sufficient quantity of data for analysis (McFadden and

Towell 1999). Here, our position is that successful cases are also needed in order to adequately identify conditions that distinguish success from failure.

Component 3: Precursor data analysis and modeling

To complement deterministic event analysis and large-volume data collection, industries have used diagnostic and predictive statistics to measure the significance of each precursor. Such data also provide further insight regarding latent patterns. In order to make probabilistic approaches a legitimate option, a large, homogeneous data set is required. As precursor programs mature and more precursor data are obtained, the use of probabilistic modeling techniques becomes a viable option (IAEA, 2004).

The primary advantage of probabilistic risk analysis is that it allows for personnel to quantitatively and objectively prioritize safety concerns. For instance, the US Nuclear Regulatory Commission (NRC) has established a threshold that a safety event can only be considered a precursor if the probability of core reactor damage is greater than one in a million (IAEA, 2004; Johnson and Rasmuson, 1996). With such a threshold, in the twenty years following the establishment of the ASP program, 422 precursors have been identified and recorded (Minarick, 1990). In contrast, British Airways used data reduction techniques and probabilistic risk analysis performed by Stephans and Talso (1997) to empirically establish the *Rule of 3*, which triggers a warning if three or more precursors are present for any given flight.

Component 4: Precursor alert systems and continuous improvement

After precursors have been identified, programs have been established to facilitate the collection of large volumes of reliable data, and advanced statistics have been used to make valid predictions, the precursor analysis system must alert the workforce when danger is detected. An effective alert system must include surveillance and quick notification of work teams when precursors are identified and conditions are ripe for a high-impact event. The ASAP, for example, mandates corrective actions to prevent precursor occurrence, which is enforced by the Federal Aviation Administration for all program members (DOT FAA/AFS-230, 2002). Similarly, the NRC ensures compliance with new regulations as precursors are uncovered (Johnson and Rasmuson, 1996) and surveillance systems are used to alert workers of precursors that could lead to reactor core damage (IAEA, 2004). NASA and the medical industry also heavily rely on surveillance alarm systems to alert workers of potential events (Bates et al., 1999; NASA/Sp-2011-3423, 2011). The primary advantage of surveillance systems is they can be built into work-flow processes very easily such as the Computerized Physician Order Entry (CPOE) system used for ordering prescription drugs in hospitals (National Academy of Engineering, 2004).

In addition to an alert system, the precursor analysis program must be built such that it can be continuously improved. Specifically, as more data become available and new factors are identified, data collection tools and predictive models must be updated. Continuous improvement is especially important in prediction because initial models typically have low skill but have the potential to improve greatly as models are tested and databases grow in quantity and improve in quality (Minarick, 1990; NASA/Sp-2011-3423, 2011). In order to create an analysis program that is fluid, a feedback loop must be established. An effective feedback system typically involves a reporting system that is regularly updated as new information is discovered or as characteristics of the work change (ETSON, 2013; McFadden and Towell, 1999; NASA/Sp-2011-3423, 2011). The industry must be committed to updating predictive models, which is not a time-consuming task, but requires knowledge in advanced statistics. The industry must also pay close attention to patterns in prediction error and seek to learn from the cause of errors.

Precursor analysis in construction

Precursor analysis for construction safety is a relatively new field of inquiry. Inspired by the nuclear industry's ASP program, Wu et al. (2010) developed a conceptual precursor model to analyze historical data to predict potential accident scenarios. Building upon the work of Cambraia et al. (2010), the Wu, et al. (2010) model closely follows Heinrich's (1931) Safety Pyramid theory by describing near-miss incidents

as the base of the pyramid with predictive validity for high-impact events. In their theory, preventing near misses is a pre-requisite to preventing high-impact events. Wu et al. (2010) use Reason's (1990) *Swiss Cheese* accident causation model and Gibb et al.'s, (2006) causation model as frameworks to classify potential precursors as either immediate, shaping, or originating factors depending on their proximity to the event. Wu et al. (2010) then defined a list of precursors specifically for scaffolding operations such as failing to wear fall protection and missing scaffolding boards. The precursors include immediate, and sometimes latent, conditions related to scaffolding work.

The research in precursor analysis for construction safety has yet to reach maturity as there is not a codified list of causal factors based on comprehensive deterministic event analysis, predictive models have yet to be created or validated, and no alert and continuous improvement systems have been established. Wu, et al., (2010) has presented a conceptual plan to implement precursor analysis in construction that shares similarities to the suggested approach presented in this study. However, the plan remains theoretical without confirmation or robust empirical validation.

Phase 3 Background: Testing the protocol through an experiment, and independent validation

Because phase 3 and Phase 4 both involve predicting safety outcomes, the background sections were combined for these phases. Please see the background for Phase 4 below.

Phase 4 Background: Creating and validating an objective predictive model

The present study builds on existing theory of construction safety risk analysis, safety leading indicators, and predictive analytics for safety. We used knowledge in these domains to elucidate our points of departure and the salient knowledge contributions. Since the companion paper presents a thorough literature review on the topic of precursor analysis, this topic is not covered in detail here. However, the salient characteristics of precursor analysis are provided below as an induction for the reader:

- The investigation is performed for one specific crew for one specific work period;
- Most precursors relate to human factors;
- Precursors are anomalous rather than routine conditions;
- Data are collected via a conversation with workers prior to work;
- Precursors are not affected by but not necessarily defined by the trade or type of work performed;
- The process is applied to work situations with the potential to severely injury or kill workers; and
- The method distinguishes success from failure using predictions made via probabilistic methods.

In the subsequent discussion, we note how these characteristics define a new approach to predictive safety that deviates from existing methods.

Safety risk analysis

Risk is generally defined as a potential event that results in an outcome that is different from what is planned. In the context of safety, the unwanted outcome is a potential injury, illness, or fatality. Researchers have consistently modeled safety risk as the product of injury frequency and injury severity as it relates to a specific unit of analysis such as a trade, activity, or working environment (Baradan and Usmen, 2006; Hallowell and Gambatese, 2009; Jannadi and Almishari, 2003). Diverse units have been analyzed using a wide range of data sources.

Units of analysis

When modeling risk, one must select a unit of analysis, defined here as the 'who' or 'what' being studied. Early safety risk studies focused on the risk of different injury types for construction trades (Brauer 1994; Hinze et al. 2005; Barandan and Usmen 2006). These risks were quantifiable because high-level data were available from the Bureau of Labor Statistics (BLS). Although analysis of safety risk for trades allowed for comparisons and prioritization, this approach was limited in terms of its application to a particular work scenario. Following these initial safety risk studies, the units of analysis have become increasingly granular

and specific. For example, Sun et al. (2008) studied specific risks associated with crane operations; Hallowell and Gambatse (2009) quantified risks for specific formwork construction activities; Esmaeili and Hallowell (2012) quantified risks for highway construction tasks; and Wu et al. (2012) studied the risks of tools involved in struck-by injuries. Very recently, Esmaeili et al. (2015) and Tixier et al. (2016) identified and validated fundamental units of analysis for construction work such as uneven work surface, work at height, ladder, and so forth. They used fundamental attributes in an effort to build a robust and generalizable method of safety risk quantification.

Although the units of analysis vary considerably across studies, all focused on physical characteristics of the work environment and associated activities. Existing risk analyses do not include human factors such as distraction, productivity pressure, and fatigue. Thus, a major way that precursor analysis deviates from safety risk analysis is via a focus on the characteristics and conditions of the workers that are present just before or during work.

Data sources

The sources of data for safety risk analysis are inherently related to the units of analysis. Risk analysts modeling general trades (e.g., Barandan and Usmen 2006) and injury types (Huang and Hinze 2003) used Bureau of Labor Statistics and Occupational Safety and Health Administration reports. The benefit is that these data are empirical; however, they are very limited in application. Alternatively, researchers who studied specific work activities (Everett 1999; Jannandi and Almishari 2003; Hallowell and Gambatse 2009) used subjective ratings of risks based on the opinions of experts because of a lack of empirical data. Unfortunately, the practical applicability gained from these specific analyses was accompanied by the judgment-based biases that plague probability assessments. To jointly address these limitations, Desvignes (2014), Esmaeili et al. (2015a, 2015b); Prades, (2014), and Tixier et al. (2016) all modeled risk of fundamental attributes using empirical data extracted from injury reports. These studies leveraged large data sources and content analysis methods such as natural language processing to extract meaningful data from injury reports.

The principal characteristics of all safety risk analysis data sources are that they are collected retrospectively from past events; include all severity levels; and include only data related to failures (i.e., injuries, fatalities, or near miss events). Precursor analysis is different because the data are collected prior to work being performed via a conversation with workers; the protocol focuses only on high-impact events (fatalities and disabling injuries); and both successes (no event) and failures (event) are included for data analysis.

Safety leading indicators

Safety leading indicators are the measures of system conditions that provide a forecast of future performance (Salas and Hallowell 2016). Researchers have begun to study leading indicators in an effort to transition from lagging indicators of performance (i.e., total recordable injury rate) to measures that can be collected and acted upon prior to an injury (Hallowell et al. 2012). These studies invariably focus on measuring the extent to which supposed leading indicators predict future safety outcomes.

Leading indicators for construction safety can be traced back to the seminal work of Liska et al. (1993) who identified proactive safety management strategies such as pre-task planning, required project training and orientation, and near miss investigation. Jaselskis et al. (1996) validated these techniques by correlating their metrics with experience modification rates and recordable injury rates. In recent years, Hinze et al. (2013) and Salas and Hallowell (2016) measured the correlation between leading and lagging indicators with increased volume of data and more precise definitions. At present, validated leading indicators include, but are not limited to, the frequency of pre-task safety meetings, frequency of safety audits, and the frequency of client participation in safety orientation. All indicators are measures of management activity that are aggregated across a large project or program of projects.

Other industries have also developed safety leading indicators programs but they tend to differ in scope and definition. For example, in the nuclear industry Reiman and Pietikäinen (2010) classified leading indicators into two groups: monitoring leading indicators and driving indicators. They claim that monitoring indicators are the objective metrics that indicate safety performance, which are analogous with the term safety leading indicator traditionally used in construction. Alternatively, driving leading indicators include situational factors like the quality of supervisors and hazard identification ability of personnel (Reiman and Pietikäinen, 2010, 2012), which are analogous to precursors according to our epistemological position.

Although construction safety leading indicators and precursor analysis both aim to identify and correct deficiencies before events occur, precursor analysis deviates in several significant ways. First, precursor analysis focuses on the conditions of the workforce rather than management activity. Second, the data are collected and analyzed for specific construction situations (e.g., a work period) rather than aggregated across a project. Finally, precursor analysis aims to predict and prevent high-impact events rather than safety performance in general. These important differences distinguish the two complementary methods of proactive safety measurement.

Predictive analytics for safety

Safety leading indicators programs and risk analyses are predictive in nature in that they involve measures of existing conditions that correlate with future performance. However, the analytical methods are limited as they do not yield probabilities of specific outcomes for a given work scenario. Researchers have recently attempted to make such probabilistic assessments using a variety of analytical methods. We used this body of literature as inspiration for our analytical methods.

Extant models

One of the earliest probabilistic models for safety was created by Chua and Goh (2005). They used a Poisson distribution and Boolean logic to model accident data and predict performance. Despite demonstrating success in predicting injury outcomes at the project level, the predictor data were too general to produce specific assessments for tasks or environments (Esmaeili et al., 2015b). Other methods of safety prediction include development of a prediction matrix based on behavior observation (Chen and Yang, 2004); neural network analysis techniques to relate safety management systems to accident data (Goh and Chua, 2013); probabilistic forecasting of loss-of-control events via expert panel workshops (Rozenfeld, et al., 2010); and prediction of injury outcomes using fundamental attributes of the work environment (Esmaeili et al. 2015b; Tixier et al. 2016).

Although a variety of predictive models have been created for construction safety, they are not able to distinguish between events or non-events (i.e., success and failure). That is, existing models are able to predict the probability of various outcomes *if an event were to occur*. This very important distinction exists because previous models used observations, analysis, or reports of injuries. Without inclusion of observations, analysis, or reports of non-event situations, it is not possible to distinguish success from failure in predictive models (Hollangel 2014). This study presents a precursor analysis that delivers a new probabilistic forecasting method that departs from existing models and approaches. Most importantly, the models created here are based upon a dataset of both success and failure, allowing for a mathematical distinction between the two.

Methods of model building

When creating predictive models, researchers often seek large datasets with many diverse predictor variables. However, when the number of predictor variables is very large, predictive models require exponentially more input data and become unwieldy and inefficient. Thus, data reduction techniques are used to reduce an initial set of predictor variables to the most meaningful by exploring covariance. For example, Cooper and Phillips (2004); Esmaeili et al., (2015b); Fang et al. (2004) all used principal component analysis (PCA) to reduce the dimensions of their data. PCA is a valuable technique that uses an

orthogonal transformation to identify a set of linearly uncorrelated variables (i.e., principal components) that serve as a smaller set of predictors.

Once the dimensions of a safety dataset are reduced, multiple regression analysis can be used to predict multiple outcomes (Johnson, 2007; Tam and Fung, 1998) or generalized linear models (GLM) with logit transformation are used to predict a binary outcome (Esmaeili et al., 2015b; Fang, et al., 2006). Both forms of analysis regress the independent variables against the dependent variable(s) to identify the predictors with the most explanatory power. Once a predictive model has been created with the base dataset, the models are generally validated against an independent dataset. This is a critical step in all predictive analytics because it explains the extent to which the predictive models are able to correctly predict the outcome of new observations better than random. We use this combination of PCA, GLM, and independent validation in our methods.

4. HYPOTHESES AND RESEARCH QUESTIONS

In this research we explored several hypotheses and research objectives. When we were able to collect empirical data or run structured experiments, we frame the inquiry in the form of a formal hypothesis. Alternatively, when we performed exploratory or developmental work, the research is framed in the form of research questions.

Phase 1 Hypotheses: Assessing the extent to which energy magnitude predicts injury severity

Hypothesis 1: Energy magnitude measureable prior to an incident does not predicts variability in injury severity

Hypothesis 2: Energy intensity measureable prior to an incident does not predict variability in injury severity

Phase 2 Questions: Identifying precursors and creating a precursor analysis protocol

Question 1: What are the causes of past construction fatalities?

Question 2: What questions may be asked of workers to assess whether the causal factors are present?

Phase 3 Hypothesis: Testing the protocol through an experiment, and independent validation

Hypothesis 3: Individuals of all levels of expertise can use the precursor analysis protocol (only leading information) and their judgment to correctly predict the outcomes of actual cases with over 80% skill

Phase 4 Hypothesis: Creating and validating an objective predictive model

Hypothesis 4: The probability of fatal and disabling events can be predicted by a small number of precursors that can be identified prior to an incident.

5. RESEARCH METHODS

Phase 1 Methods: Assessing the extent to which energy magnitude predicts injury severity

In this study, motion and gravity hazards were analyzed. We limited our scope to these energy sources because they contribute to three of the four leading classified types of construction injuries (BLS 2015). Additionally, these two energy sources proved to be readily available for robust statistical hypothesis testing. As shown in equations 4 and 5, we needed the following independent variables to compute energy magnitude and energy intensity: the weight, height, speed, and contact area of an object or person possessing the energy. These variables are leading in that they are observable prior to an incident and objectively measurable.

Although we hypothesize that energy intensity has stronger predictive validity than energy magnitude, we recognize that, in practice, the probabilistic assessments of contact area are abstract and unrealistic. Energy magnitude, however, only requires the observation of height, weight, and velocity, which are easily estimated. Table 1 illustrates several case examples of the data needed to compute energy magnitude and energy intensity from gravity and motion.

$$\text{Energy Magnitude} = mgh = \frac{1}{2}mv^2 \quad \text{Equation 4}$$

$$\text{Energy Intensity} = \frac{\text{Energy Magnitude}}{\text{Contact Area}} = \frac{mgh}{A} = \frac{\frac{1}{2}mv^2}{A} \quad \text{Equation 5}$$

where m = hazard mass; g =gravitational acceleration constant; h = height of the hazard; v = velocity of hazard; and A = contact area (i.e., “sharpness” of hazard)

Table 1: Proposed energy-based safety risk analysis computational framework with examples

	Input variables				Output Variables	
	Weight (kg)	Height (m)	Speed (m/s)	Contact Area (cm ²)	Energy Magnitude (joules)	Energy Intensity ($\frac{\text{Joules}}{\text{cm}^2}$)
Computation	A	B	C	D	$A * B$ or $\frac{1}{2} * \left(\frac{A}{g}\right) * C^2$	$\frac{(A * B)}{D}$ or $\left(\frac{1}{2} * \left(\frac{A}{g}\right) * C^2\right) / D$
Tape measure dropped from 50 stories	0.5	152.4	-	12.9	747.5	57.9
Bag of cement mix dropped from 1 story	22.7	3.0	-	232.3	668.1	2.88
Tape measure dropped from 1 story	0.5	3.0	-	12.9	14.7	1.14
Concrete chisel dropped from 1 story	0.5	3.0	-	0.65	14.7	22.6
Nail fired from pneumatic nail gun	0.001	-	30.5	0.008	0.47	58.14

The proposed hypotheses were tested by investigating the relationship between the severity of worker injury and the characteristics of the energy present before the injury was sustained. Energy magnitude and energy intensity were computed from information contained in detailed accounts of past injuries. The severity of each injury was entered into an event management system and classified according to the guidance provided by Hallowell and Gambatese (2009b). The injury severity classifications are listed in Table 2 along with their respective definitions.

Table 2: Injury classification system used during data analysis

Classification	Definition	Reports Analyzed
1st Aid	Any treatment of minor scratches, cuts, burns, etc. where the worker is able to return to work the same day following first-aid treatment.	55
Medical Case	Any work-related injury or illness requiring medical care or treatment beyond first-aid where the worker is able to return to work the following day.	113
Lost Work Time & Permanent Disability	Any work-related injury or illness that prevents the worker from returning to work the following day.	280
Fatality	Any work-related injury or illness that results in permanent aliment.	57

Injury reports were obtained from three databases: (1) donation by a consortium of 281 private construction organizations over a three year span; (2) the Fatality Assessment and Control Evaluation (FACE) reports provided by the National Institute for Occupational Safety and Health (NIOSH); (3) and the Worker's Compensation Board of British Columbia (WorkSafeBC). In total, 505 reports of injuries related to gravity and motion energy sources were obtained. The distribution of injury reports within each injury classification can be viewed in Table 2.

Processing the injury report data first involved decoupling the independent predicting variables, *hazard energy magnitude* and *hazard energy intensity*, from the dependent variable, *injury severity*. Using only descriptions contained within the written injury reports, a team of three engineering students with training in Newtonian energy principles computed hazard energy magnitude and hazard energy intensity by estimating height, weight, speed, and physical characteristics of the equipment, tools, material, and workers. Importantly, the estimation of the independent variables was performed without knowing the outcome of each case. Then, statistics were applied in order to test our hypotheses. Figure 4 illustrates the described research process.

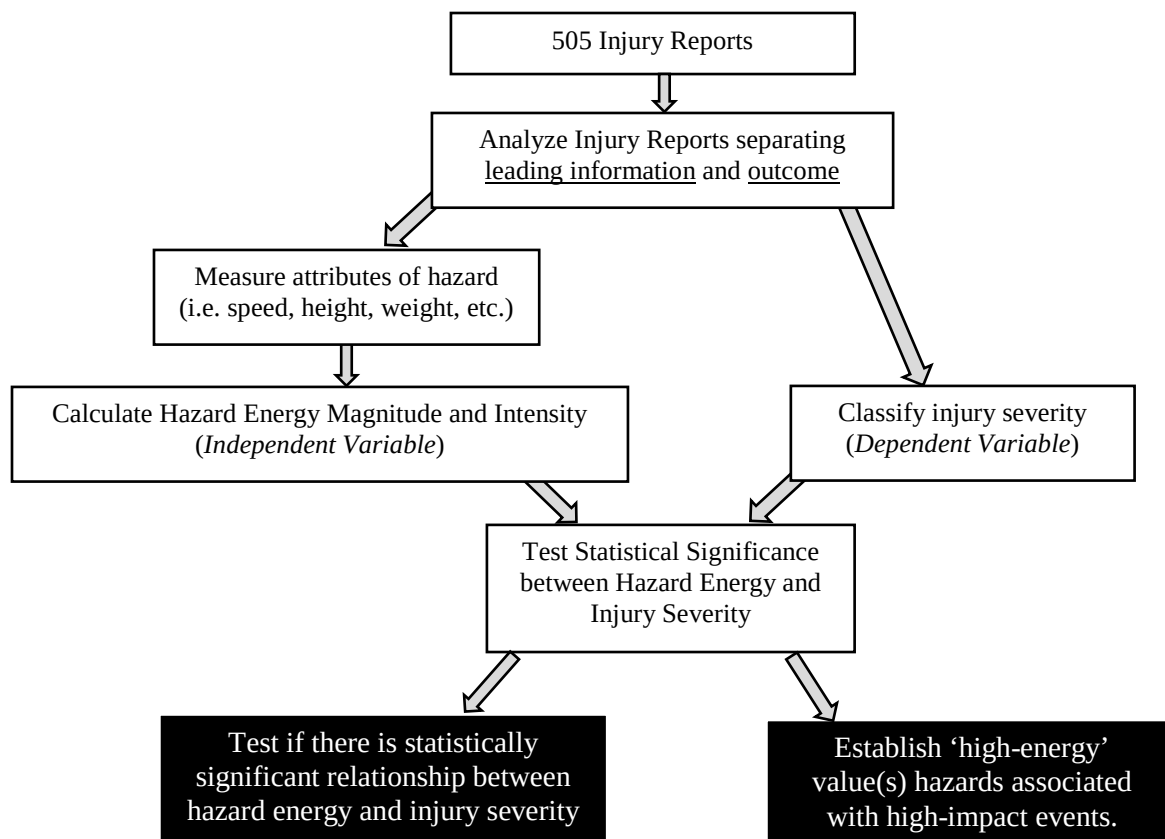


Figure 4: Diagram of research and analysis process

Phase 2 Methods: Identifying precursors and creating a precursor analysis protocol

The goal of this research was to create and validate a method for predicting the potential for construction worker fatalities through a formal precursor analysis. In other words, we aimed to create a structured method for distinguishing conditions that would likely lead to success (no event) from those that indicate potential failure (near miss, injury, or fatality). Since this was the first attempt at a data-driven method for precursor analysis, our research required a sequence of steps across two main phases. As shown in Figure 1, in the first phase we aimed to create a comprehensive catalog of potential precursors by triangulating results from a variety of sources. Although there are many studies of causal factors of construction fatalities, the potential precursors are highly dispersed in the literature. Once a catalog was established, each potential precursor was converted into a question to ask project personnel. Then, the protocol was used to collect cases of successful work, near misses, and high-impact events (fatal or disabling injuries). Once cases were collected, a series of iterative experiments were conducted to measure the extent to which individuals could use the precursor protocol and their judgment to distinguish between success and failure. It is important to note that we focused on using human judgment to make assessments because we did not yet have access to the volume of data needed to make objective predictions.

This project was completed by a Construction Industry Institute research team. The team included 14 industry experts and two academic researchers. The industry professionals represented both client and contractor organizations, all of which have sophisticated safety programs. The expertise of the research team was an important asset to enhance the validity, reliability, and practicality of the work. The expert team averaged 22 years of experience in construction safety; 10 team members possessed bachelor degrees in construction safety or related field; and four individuals had received master's degrees specifically in construction or occupational safety. Furthermore, every team member had earned at least two professional