Severity of U.S. Construction Worker Injuries, 2015-2017

David Passmore 1*, Chungil Chae 2, Victoria Borkovskaya 3, Rose Baker 4 and Jeong-Ha Yim 5

1 Penn State University, Workforce Education and Development, 305D Keller, University Park, Pennsylvania, USA 16802; 2 Penn State University, Applied Cognitive Science Lab, E365 IST Building, University Park, Pennsylvania, USA 16802; 3 Moscow State University of Civil Engineering, Yaroslavskoe shosse, 26, Moscow, 129337, Russia; 4 University of North Texas, Learning Technologies, Discovery Park, G150, 3940 North Elm Street, Denton, Texas USA 76207; 5 University of Georgia, Lifelong Education, Administration, and Policy, 850 College Station Road, Athens, Georgia, USA 30602

Abstract. Workers in the U.S. construction industry experience workplace hazards that can lead to work-related injuries that sometimes are fatal. Reported in this paper is a case-control study of risks factors associated with 4,845 injured workers and their work environments that led to fatal rather than nonfatal injuries during 2015-2017. These injury data originally were assembled from information collected by the U.S. Bureau of Labor Statistics that were used in a machine learning competition, but were repurposed for this secondary analysis of injury risks. Sixty-one percent of workers recorded in this dataset were injured fatally. Multiple logistic regression was applied to model the probability of a fatal injury as a function of the nature of the injury, part of body injured, human factors involved, whether the injured worker was carrying out a regularly assigned task at the time of the injury, and the manner in which the injury was inflicted. Related positively, relative to benchmarks, to the probability of a fatality injury were: falls and strikes; electrocution; asphyxiation and drowning; injury to the head and neck; and working at a task not regularly assigned. Negatively related to the probability of a fatal injury were: chemical/temperature burns; amputation and crushing; fractures and dislocations; injuries to fingers, hands, wrists, and other extremities; and falls from an elevation or to the same level, although this last negative relationship is anomalous in the light of independent research findings. Findings of this study do not necessarily culpable causes of work-related death. Rather these findings identify risk factors that might prove fruitful for further analysis of the incidence, severity, and costs of construction injuries.

* Corresponding author: passmore.david@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

1.1. Work-related injuries of construction workers in the U.S.

The construction industry in the U.S. employed approximately 7.5 million workers during the last quarter of 2018. In 2017, construction workers suffered 1,013 work-related fatalities and 3.1 cases of injury and illness per 100 full-time equivalent workers [1]. In 2015, the fatality rate in construction was almost three times higher than the average for all industries, while the nonfatal injury rate was almost 50% higher than the average [2]. Immigrant workers have been at especially high-risk of injury as have been older workers [3], whose injuries exact relatively high costs [4]. The demand for construction workers often is seasonal and waxes and wanes with the availability of project financing. Not surprisingly, the number of construction workers injured fluctuates seasonally [6] and over the business cycle [6, 7].

1.2. Purpose of this study

We examine the association between the severity of a construction worker’s injury and the work and worker characteristics of the injury. Specifically, estimated in this study is the relative frequentist probability of a fatal injury, rather than a nonfatal injury, of construction workers occurring between 2015 and 2017 as a logistic function of the following risk factors [each defined in more detail in 8]:

- **Nature of the injury** — The physical characteristics of an injury such as cuts and lacerations, fractures, sprains and strains, or electrocution.
- **Part of the body injured** — The part of body injured, such as a finger, arm, back, or body system.
- **Human factors associated** — Aspects of the injury that are linked to a worker’s use of defective or inappropriate equipment or flawed processes, whether by inattention, misjudgment, or deliberate action.
- **Task familiarity** — An indication of whether the worker was performing either a regularly-assigned or an unfamiliar task when the injury occurred.
- **Injury event** — The manner in which an injury was inflicted, such as caught in running equipment; slips, trips, falls, overexertion; or contact with electrical current

Task familiarity is a risk factor that exists before an injury occurs. Several risk factors are part of the causal chain that leads to an injury (i.e., human factors, injury event). Some risks of a fatality — nature of the injury and part of body injured — are direct consequences of the injury. Also, some fatality risk factors are clustered. For instance, an electrocution (nature of the injury) does not occur without contact with electrical current (the type of injury event). The time-ordering and interdependence of these risk factors are not captured in the additive model we estimate of the relationship between the severity of an injury and a variety of risk factors surrounding the injury.

2. Methods

2.1. Data

Data examined include person-level records summarizing 2,964 fatal and 1,881 nonfatal injuries to workers employed in the construction sector of the U.S. economy during 2015-2017. These injuries were reported to the U.S. Bureau of Labor Statistics through the Injuries, Illnesses, and Fatalities (IIF) program [9]. The construction sector comprises establishments engaged in the erection of
buildings or in engineering projects (e.g., highways, dams, or utility systems). Construction work done may include site preparation; new work; additions; alterations; or maintenance and repairs [10].

Injury records included in our study are selected from a complete count of work-related fatal injuries conducted through the Bureau’s Census of Fatal Occupational Injuries [11] and data collected by the Survey of Occupational Injuries and Illnesses [12], an establishment-based survey designed to estimate population-based counts and incidence rates of occupational injuries and illnesses. Injuries and illnesses are recorded if they result in loss of consciousness, days away from work, restricted job activity, job transfer, or medical treatment beyond first aid.

In this dataset, an occupational injury is any injury resulting from a single instantaneous exposure to agents of energy in a work environment [13]. In contrast, an occupational illness involves a longer latency between exposure to risks in employment and the onset of symptoms of a diagnosis, an abnormal condition, or a disorder (e.g., as in chronic exposure to ionizing radiation, pneumoconiosis, contact dermatitis, or chronic tenosynovitis). IIF records of occupational illnesses are excluded from this research, leaving injury records only. Policies and methods for accounting for occupational illnesses and injuries vary remarkably among countries [14], a limitation to the generalization of these findings beyond the U.S. construction industry.

These 4,845 records of injuries were downloaded from a public data repository [15] that provided grist for data scientists in an academic competition in machine learning. This dataset focused risk factors that differentiate fatal from nonfatal injuries in the tradition of case-control studies in epidemiology [16]. In such studies, similar numbers of cases diagnosed with a health condition and controls without the condition are compiled, even if the number of cases in the population is sparse compared with the density of controls. Sample sizes and designs in case-control studies are specified to minimize Type I and Type II statistical errors as well as to ensure adequate information about variations in risk factors [17]. As a consequence, our dataset is not suitable for estimating the incidence or degree of severity of injuries in the population of U.S. construction workers because the dataset does not a complete census of the population or a representative sample of injuries in this population.

2.2. Variables

2.2.1. Dependent

The dependent variable is the degree of severity of an injury, which is a binary variable formed by two mutually exclusive categories of viability after an injury, “fatal” and “nonfatal.” A fatal injury is coded “1”, and a nonfatal injury is coded “0” in the statistical analysis performed in this study.

2.2.2. Independent

Seventeen categorical, binary-coded (“1”, “0”) independent variables were created by regrouping and recoding the original data codes for the following five elements of the IIF data: (1) nature of the injury (seven categorical variables were created); (2) part of the body injured (two created); (3) human factors involved in the injury (two created); (4) task familiarity (one created); and (5) injury event (five created).

2.2.3. Analysis

Because the dependent variable contains two categories (“fatal” and “nonfatal”), binary logistic regression [18] was selected to estimate the probability of a fatal injury, \( P\{Y = \text{"fatal"} \} \), given the vector of 17 independent variables, \( X \), or \( P\{Y = \text{"fatal"} \mid X \} \). The range of a relative frequentist probability, \( P \), is fixed by definition as 0 to +1. A purely linear model in \( Y \) and \( X \) would allow \( P \) to exceed +1 or fall below 0. The logistic function maps the desired 0 to +1 range, where \( y = f(x) = \frac{x}{1 + x} \).

The logistic model fit to our data, then, is \( P\{Y = \text{"fatal"} \mid X \} = \frac{1 + \exp(-X \cdot \beta)}{1 + \exp(-X \cdot \beta)} \), where \( \beta \) is the slope of the log linear relationship between \( P\{Y = \text{"fatal"} \} \) and \( X \). However, interpreting \( \beta \) is cumbersome because it is an exponential. Most analysts report the
natural log of \( \log \) which is expressed as an “odds ratio” representing the ratio between the odds, in our case, that a fatality occurred and the odds that a nonfatal injury occurred, given \( X \).

The range of the odds ratio is 0 to \(+\infty\). An odds ratio = 1 indicates that a unit change in \( X \) is not related to the odds of occurrence of a fatality. An odds ratio > 1 indicates that a unit change in \( X \) increases the odds of the occurrence of a fatality. For instance, an odds ratio of 2 would reveal that a fatality is 2 times more likely to occur than a nonfatal injury, given a unit change in \( X \). However, an odds ratio < 1 is difficult to explain intuitively and can mislead [19]. Therefore, we decided to explain the results our analysis without resorting to interpretation of odds ratios.

A plot of the logistic function resembles a “lazy S,” with the slope of the plotted line depending on the specific intersections of \( P\{Y = \text{“fatal”}\} \) and \( X \). To interpret our findings from the binary logistic regression, we evaluate the first derivative of the logistic function at the mean value of \( Y \) (estimated, in our case, by the proportion of fatal injuries, or 2,964 fatalities \( ÷ \) 4,845 total = 0.6118). The first derivative of the logistic function is [20]:

\[
\frac{((\exp(-X\beta))}{(1 + \exp(-X\beta))^2} \cdot \beta
\]

Note that the expression to the left of the multiplication dot reduces to nothing more than \( P\{Y = \text{“fatal”}\} \cdot 1 - P\{Y = \text{“fatal”}\}, \) or \( PQ\beta \), where \( Q \) is \( 1 - P \).

3. Results

Displayed in Table 1 are definitions, means, and standard deviations for each of the 17 independent variables recoded from the original five data elements. Point and interval estimates of \( PQ \mid \) are shown in the column situated furthest to the right in Table 1, with the product of \( P = 0.6118 \) and \( Q = 0.3882 \) multiplied by each element in the vector of \( \mid \) estimated in the binary logistic regression. Also provided in a footnote to Table 1 is the logistic regression equation estimated for \( P\{Y = \text{“fatal”} \mid X\} = [1 + \exp(-X\beta)]^{-1} \) along with associated measures of the goodness-of-fit of the logistic regression model to the data.

Estimates of \( PQ \mid \) displayed in red in Table 1 indicate a negative relationship of a risk factor with \( P\{Y = \text{“fatal”}\} \), while estimates shown in green are positively associated. All other estimates of \( PQ \mid \) visible in black-colored type exhibit 95% confidence intervals around \( PQ \mid \) that include the value of 0.0 in the range between the upper and lower real limits of the interval, which means that the point estimate of \( PQ \mid \) are not distinguished statistically as different than zero.

3.1. Nature of injury

On one hand, chemical and high/low temperature burns, amputation and crushing, and fractures and dislocations are less likely to result in death than bruises, lacerations, contusions, or punctures. Deaths, on the other hand, are more likely from falls and strikes, electrocution, and asphyxiation and drowning than from bruises, lacerations, contusions, or punctures. Construction is estimated to greatest proportion of workplace electrical injuries globally [21].
3.2. Part of body injured

A higher probability of death results from injuries to the head and neck than from injuries to the torso of the body (an area independent of the head, neck and limbs). Injuries to fingers, hands, wrists, and other extremities are less likely to lead to death rather than injuries to the torso.

3.3. Human factors

Indications of use of defective or inappropriate equipment or occurrence of incidents involving misjudgment, inattention, or deliberate action do not help to discriminate fatal from nonfatal injuries. Other research [22] has identified “inattentional blindness” — when workers focused on their main task ignore other dangers in their environment — as a persistent safety problem in the construction workplace.

3.4. Task experience

Injuries resulting in a fatality are more likely when workers perform tasks not regularly assigned to them.

Table 1. Probability of fatality among injured construction workers by nature of injury, part of body injured, human factors, task experience, and event type, 2015-2017 (n = 4,845).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (Standard Deviation)</th>
<th>$PQ\beta$ (95% confidence interval)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatality (dependent variable)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatal</td>
<td>0.61 (0.49) = $P$</td>
<td>—</td>
</tr>
<tr>
<td>Nonfatal</td>
<td>0.39 (0.49) = $Q$</td>
<td>$rc$ = reference category</td>
</tr>
<tr>
<td>Risk Factor (independent variable)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Nature of Injury</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Chemical/temperature burns</td>
<td>0.04 (0.19)</td>
<td>-0.17 (-0.28; -0.07)</td>
</tr>
<tr>
<td>b. Amputation and crushing</td>
<td>0.11 (0.32)</td>
<td>-0.38 (-0.53; -0.24)</td>
</tr>
<tr>
<td>c. Falls and strikes</td>
<td>0.35 (0.48)</td>
<td>0.45 (0.39; 0.50)</td>
</tr>
<tr>
<td>d. Electrocution</td>
<td>0.04 (0.20)</td>
<td>0.58 (0.35; 0.82)</td>
</tr>
<tr>
<td>e. Fractures and dislocations</td>
<td>0.19 (0.39)</td>
<td>-0.11 (-0.17; -0.06)</td>
</tr>
<tr>
<td>f. Head trauma</td>
<td>0.07 (0.25)</td>
<td>0.6 (-0.01; 0.14)</td>
</tr>
<tr>
<td>g. Asphyxiation and drowning</td>
<td>0.04 (0.21)</td>
<td>0.71 (0.55; 0.89)</td>
</tr>
<tr>
<td>h. Bruises, lacerations, contusions, and punctures</td>
<td>0.39 (0.49)</td>
<td>$rc$</td>
</tr>
<tr>
<td>2. Part of Body Injured</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Head and neck</td>
<td>0.26 (0.44)</td>
<td>0.28 (0.23; 0.32)</td>
</tr>
<tr>
<td>b. Fingers, hands, wrists, and other extremities</td>
<td>0.50 (0.50)</td>
<td>-0.71 (-0.82; -0.61)</td>
</tr>
<tr>
<td>c. Torso</td>
<td>0.24 (0.43)</td>
<td>$rc$</td>
</tr>
</tbody>
</table>
### 3. Human Factors

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>( \beta ) (SE)</th>
<th>( \beta ) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Use of defective, malfunctioning, insufficient, and inappropriate equipment</td>
<td>0.09 (0.29)</td>
<td>-0.06 (-0.13; 0.01)</td>
</tr>
<tr>
<td>b. Lapses in worker judgment</td>
<td>0.53 (0.50)</td>
<td>-0.02 (-0.06; 0.01)</td>
</tr>
<tr>
<td>c. Other</td>
<td>0.38 (0.49)</td>
<td>\textit{rc}</td>
</tr>
</tbody>
</table>

### 4. Task Experience

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>( \beta ) (SE)</th>
<th>( \beta ) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Working on a task not regularly assigned</td>
<td>0.37 (0.48)</td>
<td>\textit{0.07 (0.03; 0.11)}</td>
</tr>
<tr>
<td>b. Working on a regularly assigned task</td>
<td>0.63 (0.48)</td>
<td>\textit{rc}</td>
</tr>
</tbody>
</table>

### 5. Event Type

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>( \beta ) (SE)</th>
<th>( \beta ) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Caught in or between…</td>
<td>0.23 (0.42)</td>
<td>-0.02 (-0.10; 0.06)</td>
</tr>
<tr>
<td>b. Fall from elevation or same level</td>
<td>0.27 (0.44)</td>
<td>\textit{-0.10 (-0.18; -0.03)}</td>
</tr>
<tr>
<td>c. Struck by or against…</td>
<td>0.26 (0.44)</td>
<td>0.06 (-0.02; 0.13)</td>
</tr>
<tr>
<td>d. Respiratory or cardiac failure</td>
<td>0.05 (0.23)</td>
<td>0.05 (-0.07; 0.16)</td>
</tr>
<tr>
<td>e. Electric shock</td>
<td>0.04 (0.20)</td>
<td>0.16 (-0.06; 0.38)</td>
</tr>
<tr>
<td>f. Other</td>
<td>0.14 (0.34)</td>
<td>\textit{rc}</td>
</tr>
</tbody>
</table>

**Source:** Analysis of OSHA accident and injury data: Injury records for 2015-2017 [8]

\[ \beta \] values are the coefficients in the following logistic regression equation (variable numbers and letters identify numbered and lettered risk factor variable names from the table):

\[
P\{Y = \text{“fatal”} \mid X\} = [1 + \exp(-(0.01(\text{intercept}) + (-0.17[1.a]) + (-0.38[1.b]) + (0.45[1.c] + (0.58[1.d]) + (-0.11[1.e]) + (0.06[1.f]) + (0.71[1.g]) + (0.28[2.a]) + (-0.71[2.b]) + (-0.06[3.a]) + (-0.02[3.b]) + (0.07[4.a]) + (-0.2[5.a]) + (-0.10[5.b]) + (0.06[5.c]) + (0.05[5.d]) + (0.16[5.e]))}]^{-1}, \]

-2log likelihood = 2591 with 17 df, \( p < 0.001 \), indicating that the 17 risk factors accounted for a non-zero amount of variation in the degree of injury severity. The logistic equation accounted for 88\% of the area under a receiver operating characteristic curve, demonstrating the ability of the equation to correctly discriminate between fatal and nonfatal injuries.

### 3.5. Event type

One type of event, a fall from an elevation or at the same level, less probably results in a fatality than “other” types of events. This finding seems anomalous because a fall, especially from a height, might involve considerable kinetic energy that would bluntly assault a body, which could elevate the risk of death compared with “other” events. Falls were the leading cause of death and the third leading cause of non-fatal injuries in the U.S. construction industry in 2013 [23].

Insensitivity in the IIF program to measuring and coding injury data could account for this finding. Incomplete or inadequate reports to the IIF program might result in grouping some hazardous events under the category of “other” unspecified events. Or, perhaps the IIF program does not implement processes adequate to measure or describe events completely or reliably. As a consequence, the “other” category could be nothing more than a diffuse dustbin of hazards that are not explainable otherwise.

### 4. Discussion

#### 4.1. Cautions

Our findings do not quantify the risk of death or chances of survival from work-related injuries that the population of construction workers faces. Rather, this study assesses the
risk of death among workers who already are injured. This limitation on our findings is dictated because the dataset we analyzed is not assembled from a complete census or representative sample of all construction workers, injured or not.

Information usually is unavailable about the population at risk of injury. Instead, information more often is available only about injured construction workers. And, within this limited information, risks of death are calculated for injured workers only relative to benchmark categories of risk defined by the reference category (rc entries) in Table 1.

4.2. Evidence useful for injury prevention

Our findings fit in the middle of a chain of evidence that is useful to develop incentives, strategies, and interventions designed to prevent construction worker injuries. The first stage in the chain of evidence is an assessment of the incidence of work-related injuries among construction workers. A work-related injury incidence rate equals the number of workers injured divided by the number of population of workers at risk for injury over a specified time period. We might learn, for instance, that three deaths occur for every 100 workers over a period of a year. Consideration of incidence rates helps set priorities for investment in injury prevention. Because risks probably vary by type of construction activity (e.g., highways, building structures site preparation), incidence rates are most useful if they are activity-specific. What is crucial, however, is to have information about the population at risk. Yet, in most industries, such population counts are difficult to obtain. And, construction is no different.

The evidence provided through this particular study of the severity of construction injuries points to various hazards that can multiply the consequences of a construction injury that has occurred [24]. The aim of studies in the second stage in the chain of evidence is to identify risk factors that raise or lower the severity of injuries once they occur. An example of this type of second–stage evidence from another injury domain would be an analysis of the effectiveness of seat belts in reducing the severity of injuries after an automobile collision. Seat belts do not affect whether a collision happens, but, post-collision, seat belts can mitigate the severity of injuries by reducing chances of ejection from the auto, blunt trauma from striking the steering column, or contusions or lacerations resulting from forceful impact with surfaces or protuberances in the passenger compartment of the auto. Similarly, information about factors related to the risk of severe injuries in construction can lead to better engineering and workplace designs.

Studies investigating the second stage in the chain of evidence typically rely on observations derived directly from the injury incident to identify associations between risk factors and the severity of injuries. However, the third stage of evidence necessary for designing evidence-based incentives, strategies, and interventions for injury prevention is information about the costs exacted by high incidence and severe injuries [25, 26]. Information from the entire chain of evidence helps policy makers focus on prevention designed to prevent an optimal mix of the highest incidence, highest severity, and most costly construction worker injuries [27].

4.3. A few additional research possibilities

Although the need for additional information about the epidemiology of severe injuries is evident, we suggest merely a few promising research possibilities. First, one factor that could affect the severity of injuries is firm size [28, 29, 30]. In particular, smaller firms might be less equipped than larger firms to respond to an injury once it occurs, which could delay critical medical response. Small firms also might have less diversity of experience with hazardous work situations than larger firms. Second, another research
possibility exists in extending systems-based [31, 32] and systematic approaches to analysis of hazardous work situations [33, 34].

5. Conclusions

Concluded from an analysis of the risk factors associated with the severity of 4,845 injuries in the U.S. construction industry is that:

- Fatalities are more likely than nonfatal injuries as a result of:
  - Falls and strikes, electrocution, and asphyxiation and drowning compared with bruises, lacerations, contusions, and punctures;
  - Injuries to the head and neck compared with injuries to the torso; and
  - Working on a task not regularly assigned.
- Fatalities are less likely than nonfatal injuries when:
  - Chemical/temperature burns, amputation and crushing, and fractures and dislocations occur rather than bruises, lacerations, contusions, and punctures;
  - Injuries occur to fingers, hands, wrists, and other extremities rather than to the torso; and
  - A fall from an elevation or to the same level occurs rather than “other” (unspecified) situations, although we find this outcome anomalous in the light of independent research findings.

References

22. Liao C Chiang T 2016 Reducing occupational injuries attributed to inattentual blindness in the construction industry *Safety Sci.* 89 129-137 doi:10.1018/j.ssci.2016.06.010


34. Borkovskaya V 2018 Project management risks in the sphere of housing and communal services. *Journal MATEC Web of Conferences* **251** doi:10.1051/matecconf/201825106025