Journal of Safety Research 98 (2026) 175–189

Contents lists available at ScienceDirect

Journal of Safety Research

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

Balanced approach to serious injury and fatality prevention: Exploring empirical relationships between short and long-term measures of safety performance

Arnaldo Bayona[a,1,*], Matthew R. Hallowell[b], Siddharth Bhandari[b], Yaqoob Raheemy[b]

[a] Department of Engineering Technologies, Safety, and Construction, Central Washington University, Ellensburg, WA 98926, USA
[b] Construction Safety Research Alliance, University of Colorado Boulder, Boulder, CO 80309, USA

ARTICLE INFO

Keywords:
Serious injuries and fatalities
Safety metrics
Leading indicators
Safety performance

ABSTRACT

Introduction: Safety is increasingly being defined not by the absence of harm, but by the presence of capacity to control critical hazards. However, widely used metrics like the rate of recordable incidents can obscure true exposure to conditions during work that lead to serious injuries and fatalities (SIFs), especially when these metrics are used in isolation. To address this gap, we examined whether two new short-term measures of SIFs, namely the quality of pre-job safety briefs (PJSB quality) and high-energy control assessments (HECA), are statistically related to conventional and alternative lagging indicators. *Method:* Field assessment data and injury records from 31 utility and construction organizations were analyzed to test these relationships. *Results*: The results indicate that: (1) a lower injury rate in the past is associated with stronger short-term safety performance in the future, as measured by more engaged pre-job safety briefs and reduced exposure to high-energy hazards; (2) PJSB quality is positively associated with the HECA score over short periods; and (3) a higher HECA score is associated with reduced SIF rates over extended periods. These results indicate that monitoring key safety activities and indicators of energy control can serve as short-term measures of SIF exposure that predict future outcomes. Additionally, combining traditional and contemporary metrics offers a more holistic representation of SIF exposure, one that balances leading, monitoring, and lagging indicators. *Conclusions and practical applications:* Such information may help companies transition from reactive approaches to safety measurement to a balanced approach that considers input, system monitoring, and outputs.

1. Introduction

Although there have been decades of progress in reducing nonfatal construction injuries, fatal incidents have proven far more resistant to change (BLS, 2025). For too long, safety programs have operated on the assumption that preventing minor injuries would eventually reduce fatalities, based on a presumed fixed ratio between low and high-severity incidents (Anderson & Denkl, 2010; Bayona et al., 2025; Busch et al., 2021; Krause & Murray, 2012; Martin & Black, 2015). However, this presumption is increasingly disputed by national surveillance data and studies comparing causes of various injury severity types (e.g., Bayona et al., 2024; Hallowell et al., 2021). Mounting evidence suggests that the characteristics of hazards themselves differentiate between high and low-severity outcomes (Bellamy, 2015). Specifically, serious injuries and fatalities (SIFs) are characterized as the outcomes resulting from uncontrolled high-energy releases (Manuele, 2005; Martin & Black, 2015; Hallowell et al., 2017), where existing barriers and task-specific work plans were either absent, inadequate, or transgressed (Bayona et al., 2024). Thus, practitioners are moving away from generalized safety efforts to those that target SIFs prevention (EEI, 2025). Despite this shift, companies continue to measure safety performance based on injury rates and other similar measures that are dominated by high-frequency, low-consequence events, such as nonfatal recordable injuries (Janicak, 2009). Consequently, this approach to measuring safety disproportionately focuses on lower-severity injuries due to a skewed dataset, obscuring the learnings that could enable SIF-sensitive

---

* Corresponding author.

E-mail addresses: arnaldo.bayona@cwu.edu (A. Bayona), matthew.hallowell@colorado.edu (M.R. Hallowell), siddharth.bhandari@colorado.edu (S. Bhandari), mohammadyaqoob.raheemy@colorado.edu (Y. Raheemy).

[1] Formerly affiliated with the Construction Safety Research Alliance, at the time the research was conducted.

https://doi.org/10.1016/j.jsr.2026.06.007
Received 31 October 2025; Received in revised form 12 March 2026; Accepted 18 June 2026

practices.

To address SIFs, organizations need measurement approaches that focus on conditions that are predictive of high-consequence events. Such approaches move beyond incident recordkeeping to incorporate indicators that: (1) capture high severity potential during ongoing work, (2) align SIF differentiators with contemporary safety paradigms, and (3) reflect valid correlations between the variables of interest. To put it differently, organizations need to know which safety efforts translate to long-term reductions in SIFs. Consequently, some researchers and practitioners have explored SIF-focused metrics that prioritize severity over frequency (Cooper, 2019). These include SIF precursor analysis (e.g., Krause & Murray, 2012; Martin & Black, 2015; Alexander et al., 2017; Hallowell et al., 2017), task-based risk assessments (Taubitz, 2018), and high-energy control assessments (Oguz Erkal & Hallowell, 2023), which all focus on the work conditions that could lead to SIFs. Collectively, such assessments offer the potential for a balanced and targeted approach to SIFs prevention. When used together, they may help identify where safety interventions are needed, how much is required, and how improved short-term safety capacity relates to long-term safety outcomes (Oguz Erkal & Hallowell, 2023). Although a balanced approach to safety measurement has been proposed in the literature (e.g., Lofquist, 2010; Oguz Erkal et al., 2023; Reiman & Pietikäinen, 2012), the empirical links between short-term metrics and long-term outcomes have yet to be demonstrated.

Thus, we tested the statistical relationship between a SIF-focused leading indicator, a novel monitoring indicator, and injury rates for 31 different utility and construction organizations. Specifically, we tested the hypothesis that quality of pre-job safety meetings and the extent to which high-energy hazards are controlled (i.e., two short-term measures of SIFs) have an empirical relationship with long-term SIF rates. Examining these connections is a key step towards understanding how short-term measures of safety activities and conditions translate into long-term safety performance.

2. Definitions of safety

Safety metrics are intimately tied to definitions. Without a shared definition, a metric can misrepresent what it intends to measure. Therefore, before discussing safety metrics, we must first define what we mean by “safety.” The widely cited maxim “if you cannot measure it, you cannot improve it” (Jazayeri & Dadi, 2017) rests on a deeper principle: *we cannot measure what we cannot define*. Therefore, before structuring and testing new metrics, we must first define safety and assess how well those metrics align or fail to align with that definition.

There is a longstanding lack of consensus around what safety means, which has yielded a fragmented set of definitions and misaligned metrics. Our work is positioned around Conklin’s (2019) view of “safety as the presence of defenses, not merely the absence of events” (see Vandeskog, 2024). This view aligns with contemporary theories such as High-Reliability Organizations, Resilience Engineering, Safety Differently, Safety-II, and Human and Organizational Performance, which emphasize adaptability, learning, and systemic defenses (Ball & Frerk, 2015; Conklin, 2019; Provan et al., 2020).

Across contemporary definitions, safety is consistently framed as an active capacity rooted in defenses, hazard control, and the conditions that make work go right. Accordingly, safety is described as preserving positive value (Vandeskog, 2024), enabling intended system performance while preventing harm (Provan et al., 2020), emphasizing controls rather than outcomes alone (Conklin, 2019), and reducing risk through sustained hazard control and accountability (Balderson, 2016), alongside both preventing things from going wrong (Hollnagel, 2014) and strengthening the capabilities that make things go right (Dekker, 2014).

2.1. Measuring what we can define

The intent of this review is not to anoint a single “correct” definition of safety; rather, it is to identify common themes across definitions and explain why particular metrics were selected for this study. Problems arise when safety is defined from one theoretical perspective but measured from another. For example, defining safety as the presence of defenses while relying on lagging indicators such as Total Recordable Incident Rate (TRIR) will not reflect that aim; conversely, defining safety as the absence of harm while using proactive indicators may appear successful until a serious incident occurs. Such misalignment creates a false sense of security, shifts attention toward peripheral metrics, and undermines benchmarking, since measures grounded in incompatible constructs cannot be compared with integrity. Therefore, before structuring or evaluating safety metrics, it is important to assess whether selected metrics are conceptually aligned and to build a balanced set that tells a more complete story.

Raheemy et al. (2025) proposed a framework (see Fig. 1) based on input from 518 safety professionals that uses a temporal lens to define safety in relation to time. In this model, future-oriented definitions focus on proactive actions (inputs), present-oriented definitions emphasize current system conditions (monitoring), and past-oriented definitions reflect on outcomes (outputs). These categories correspond to familiar types of safety metrics: leading indicators, monitoring indicators, and lagging indicators, respectively. While both leading and monitoring indicators can be gathered in real-time and reported frequently or any other shorter-term span, lagging indicators are long-term, as they usually require multiple months or years to reach stable statistical conclusions (Hallowell et al., 2021).

Rather than forcing a singular definition, the framework makes explicit which view is being applied and enables more precise evaluation of metric alignment. In the sections that follow, we use this structure to assess how commonly used metrics reflect—and sometimes distort—the definitions they are intended to represent.

3. Safety metrics

Construction organizations rely, to varying extents, on leading, monitoring, and lagging indicators to measure both short and long-term site safety performance (Janicak, 2009). Measurement approaches and associated dependencies vary greatly across organizations, and companies often evaluate these indicators separately rather than collectively (Lofquist, 2010). Before examining how safety indicators interrelate, it is important to examine how specific metrics align with the different conceptualizations of safety and acknowledge their inherent strengths and limitations.

3.1. Safety as outputs (lagging indicators)

Traditionally, safety has been measured only by incident rates (Raheemy et al., 2025). This view considers safety as process *outputs* and believes that the absence of injuries signifies progress (Petersen, 1998). Given their desire to reduce incident rates, many organizations continue to make business decisions using this output approach alone because these metrics are clear, simple, and consistently applied, which makes them relevant across organizations (Hallowell et al., 2021; Petersen, 1998). Despite their simplicity, lagging measures have a host of severe limitations such as:

- First, they are inherently retrospective and reactive. Their absence may create a false sense of safety as they do not reflect an organization’s capacity to prevent future incidents or reduce fatality risk (Lingard et al., 2017).
- Second, these lagging metrics are prone to underreporting, especially when tied to incentive programs (Lofquist, 2010; Lingard et al., 2017; Versteeg et al., 2019).

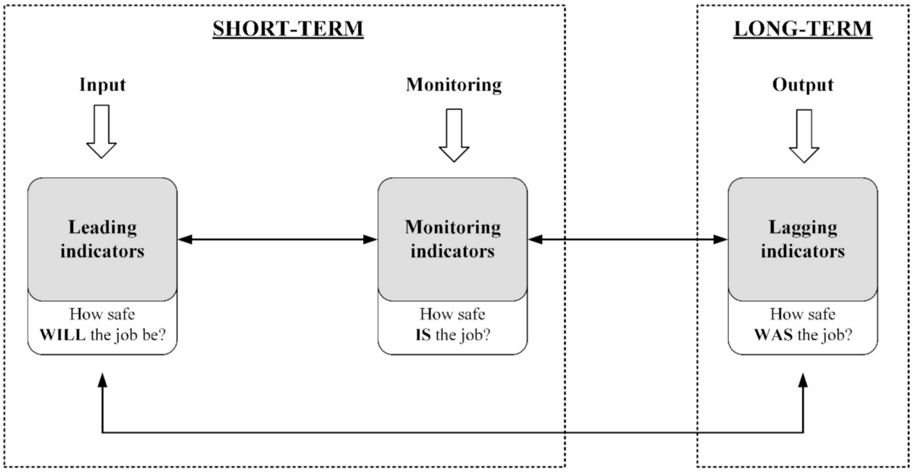

Fig. 1. Framework aligning safety metrics with contemporary definitions of safety.

- Third, to be statistically stable and meaningful, they require millions (even hundreds of millions in some cases) of worker-hours of exposure time, which for most companies takes years to accumulate (Hallowell et al., 2021; Kjellén, 2009; Oguz Erkal & Hallowell, 2023).
- Fourth, all injuries are typically weighted equally, regardless of their extremely different human consequences (Hallowell & Oguz Erkal, 2024). For example, both a fatality and a minor cut on the hand are counted as equally-weighted single incidents when calculating TRIR.
- Finally, due to their retrospective nature, many lagging indicators offer limited reliability as predictors of future safety performance (Hallowell et al., 2021; Hoonakker et al., 2005; Lofquist, 2010; Petersen, 1998).

3.2. Safety as inputs (leading indicators)

Although safety has been traditionally measured through past incidents (Hinze et al., 2013), researchers have also explored leading indicators, which are measures of safety activities that are framed as input metrics that predict future safety performance (Hallowell et al., 2013). Unlike lagging indicators, leading indicators provide an understanding of the capacity of the organization to operate safely. Measuring the activities companies perform to promote safety reflects inputs that “prioritize anticipation and preparation for future safety conditions” (Raheemy et al., 2025, p.8). Examples include the percentage of job site toolbox meetings attended by job site managers, the percentage of audited items in compliance, among others (Hinze et al., 2013).

Although companies are increasingly measuring safety through leading indicators, these practices remain overly reliant on frequency-based rather than quality-focused metrics (Oswald, 2020; Bayramova et al., 2023). This practice hinders the capacity of leading indicators to reveal weaknesses in safety systems and limits the validity of these as metrics of safety performance (Golabchi et al., 2024, p. 11). Just like lagging indicators, frequency-based leading activities are limited by how many incidents have occurred, as there is a substantial latency between program implementation and measurable impact. For instance, a low frequency of pre-job safety meetings or safety inspections does not identify risk factors such as poor hazard awareness, limited safety controls, or a lack of understanding of work procedures, unless the company has consistently experienced a poor incident record. A more robust approach shifts from counting activities to evaluating how well they improve onsite safety performance and help reduce immediate exposure to SIFs (Oswald, 2018; Oswald, 2020). For example, Bayona et al. (2024) found that SIFs are preceded by a lack of proper pre-job planning, where controls were deemed absent or inadequate.

3.3. Safety as indicators of present conditions (monitoring indicators)

Monitoring indicators are defined here as a subset of leading indicators that monitor the ability of systems to produce safe outcomes. This approach focuses on addressing safety *conditions* in the present time by checking whether workers are guarded against hazards during ongoing work (Raheemy et al., 2025). Monitoring indicators of safety performance, such as compliance audits or observations of active work, may help overcome the limitations of relying solely on long-term outcomes by measuring potential losses over the short term (Janicak, 2009, p.28) and allowing organizations to adjust their safety strategies during ongoing work. For example, they may track existing controls, worker behavior, and process safety. However, unlike leading indicators, most monitoring indicators can be largely influenced by extraneous factors that are distal to the work environment (Lingard et al., 2015; Reiman & Pietikainen, 2012). This issue has motivated researchers (e.g., Kjellén, 2023) to investigate short-term measures of available physical barriers against high-consequence events.

To support emerging approaches in SIF prevention and to align directly with the concept of safety as the presence of safeguards, Oguz Erkal and Hallowell (2023), introduced High-Energy Control Assessment (HECA). As defined by the authors, HECA is the proportion of high-energy hazards (i.e., those exceeding 1500 Joules of physical energy) with a corresponding Direct Control. Examples of high-energy hazards include falls from elevations greater than 4 feet, mobile equipment with workers on foot, and a suspended load. A Direct Control has been defined as a safeguard that is directly targeted to the high-energy source; mitigates the energy to acceptable levels where SIFs are no longer the most likely outcome; and works even if there is an unintentional error unrelated to the installation, use, and verification of the controls (Hallowell & Spencer, 2024). Examples of Direct Controls include hard physical barriers, de-energization with lock-out tag-out (LOTO), fall arrest systems, and some specialty personal protective equipment like blast suits. Although HECA is promising, its link to long-term SIF outcomes has yet to be demonstrated empirically (Oguz Erkal & Hallowell, 2023).

Each type of safety metric provides different information and has unique strengths and weaknesses. Therefore, exploring the connections among multiple metrics provides far more valuable insight than using metrics in isolation (Golabchi et al., 2024; Oguz Erkal et al., 2024). Although the connections among leading, monitoring, and lagging indicators have been theorized by several authors (e.g., Lofquist, 2010; Reiman & Pietikäinen, 2012), most empirical research has focused on the leading-lagging relationships. Furthermore, although qualitative indicators of safety have been proposed (Oswald, 2020), none are linked to other types of (safety) condition monitoring or lagging metrics.

4. Hypotheses and contribution to the body of knowledge

The leading-lagging relationship has been widely documented in literature, with studies highlighting some leading indicators, such as the frequency of pre-job safety briefs and safety inspections, as stable predictors of long-term injury rates (Alruqi & Hallowell, 2019). However, to the authors' best knowledge, no study has yet examined the empirical links among quality measures of safety activities, safety conditions, and long-term injuries. Furthermore, the empirical relationship between safety indicators integrated into the three temporal themes depicted in Fig. 1 has yet to be validated empirically. This work contributes to the body of knowledge by:

1. Examining the relationship between assessments of pre-job safety brief quality and work conditions that follow immediately.
2. Examining the connection between condition monitoring and long-term injury rates.
3. Assessing the link between both quality and condition measures and long-term performance.

This paper explores these connections by looking at one indicator example from each of the temporal themes described in Fig. 1, namely, pre-job safety brief (PJSB) quality (a leading indicator), HECA (a monitoring indicator), and long-term injury rates (a lagging indicator). This approach deviates from previous research on safety metrics by directly testing the empirical relationships among two short-term metrics directly associated with SIF conditions (PJSB quality and HECA), with long-term injury rates. To this end, the authors sought to answer the following research question:

What is the statistical relationship among PJSB quality, the HECA score, two short-term measures of SIFs, and long-term injury rates across different companies?

To answer this question, the following hypotheses were tested using a mix of linear and generalized regression models:

> *Companies with higher-quality PJSB have more Direct Controls.*
>
> *Companies with lower injury rates have better PJSB and more Direct Controls.*
>
> *Some types of firms have higher PJSB quality and more Direct Controls than others.*
>
> *Companies with better PJSB and more Direct Controls experience fewer injuries.*

By testing these hypotheses, the research will elucidate the underlying relationships that govern SIF-focused input, monitoring, and output metrics. Practically, this knowledge will support the implementation of a combined and balanced approach to safety measurement through targeted safety indicators that are most predictive of SIF risks.

5. Research methods

To test the hypotheses, a field experiment was designed and conducted on active construction sites across the United States and Canada. Multiple trained safety professionals evaluated distinct tasks representative of the work performed by utility and construction participant organizations. A multi-site and practitioner-led assessment approach provided the necessary data to explore how the company's long-term injury records correlate with the quality of pre-job safety briefs and the monitoring of site conditions. (Kjellén, 2023; Laitinen et al., 1999). A diverse and sufficiently large sample promoted a balance between external and ecological validity by mitigating the risk of losing participant companies while conducting field assessments in naturalistic settings.

To reflect the varying severity of injuries, the authors examined both broad injury rates (i.e., rates including a wide range of injury types rather than a single category) and specific lagging safety metrics (e.g., the rates of medical treatment or lost-time injuries). Two broad injury rates were included: TRIR and the severity-based lagging indicator, or SBLI. Adding SBLI to the research protocol ensured that metrics were weighed differently based on severity, unlike traditional metrics like TRIR, where injuries are weighed equally.

The overall research process is outlined in Fig. 2. This multi-phase effort involved: (1) recruiting utilities and construction companies; (2) training safety professionals within organizations in the field data collection protocol (hereafter referred to as the assessors); (3) obtaining historical injury records of varying severities for each participating company; (4) collecting field assessment data on pre-job safety briefs and HECA to obtain stable baseline measures of short-term safety and validate these data using established inter-rater reliability protocols; (5) cleaning and pre-processing the data; and (6) testing the hypotheses using linear and generalized regression models.

5.1. *Recruitment of participants and sample size required*

Participants were recruited through the networks of the Construction Safety Research Alliance (CSRA). A total of 31 companies across North America, representing utility clients and heavy civil, industrial, and commercial construction contractors, participated in the experiment. This combined sample enhanced the external validity of the results. Each participant company provided at least one assessor who collected the field assessment data for their corresponding organization. This collaboration with industry practitioners enabled the collection of vastly more data than is logistically possible with one researcher gathering data alone.

To get an adequate sample size, the authors distributed the data collection responsibilities but ensured its quality by training the assessors and conducting inter-rater reliability tests. A minimum sample size of 15 field observations for pre-job safety briefs plus 15 HECAs per company was needed to ensure that the average baseline measures were within acceptable ranges (i.e., within an 85%-confidence interval). This estimate assumed a sample standard deviation of 25% observed in previous trials (Bayona Malo, 2025), a statistical power of 80%, and a 5% alpha level. To observe a meaningful correlation between the variables, at least 26 organizations were required in the full-scale study.

5.1.1. *Training of assessors*

After recruitment, the assessors were trained and tested in the application of the PJSB quality and HECA data collection protocols (see Table 1). A single-day training workshop was delivered by the designated research personnel, using a standard training protocol for both HECA and PJSB quality. The training began with an overview of high-energy hazards and Direct Controls, followed by a description of the qualities of a high-quality pre-job safety meeting and the elements of the associated quality assessment rubric (see Table 2; CSRA Construction Safety Research Alliance. n.d.). Consistency was maintained by using standardized materials, including slides, interactive videos, talking points, and handouts. At the end of the training session, participants should have been able to:

(1) Distinguish between low and high-energy hazards and between Direct Controls and other barriers;
(2) Identify the elements of an excellent pre-job safety brief discussion;
(3) Assess the quality of pre-job safety briefs using a standardized checklist;
(4) Practice with their colleagues to ensure consistency in the application of the protocol.

To practice with their colleagues in each organization, standardized data collection protocols for HECA and PJSB quality assessment were shared and then applied through several practice scenarios. These scenarios were produced and validated by the Edison Electric Institute (EEI

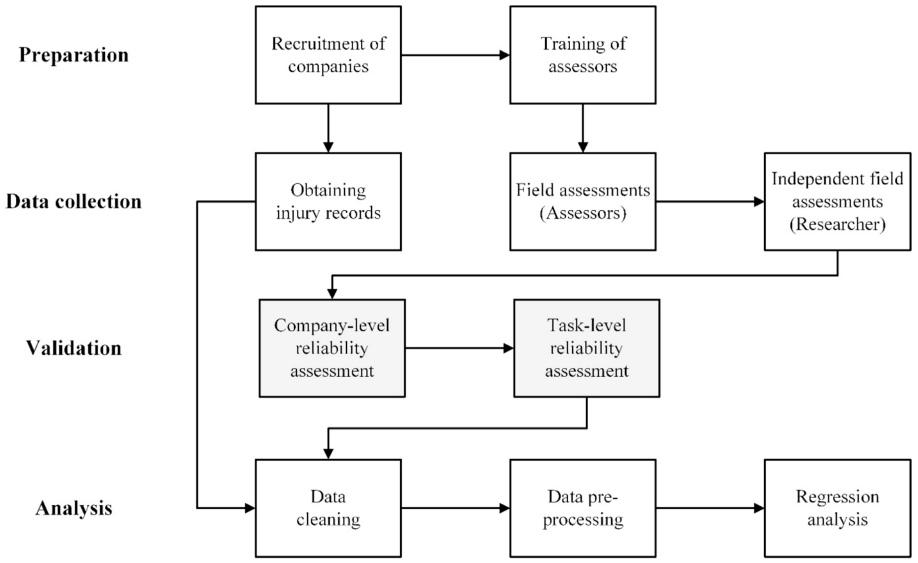

Fig. 2. Research process.

Table 1
Field data collection instruments.

Variable	Description of instrument	Example statements
PJSB quality score	Checklist with binary assessment (yes or no) of 15 statements related to the task, hazards, and controls discussed during a pre-job safety meeting (See Table 2).	Controls for each identified hazard were identified and discussed (Yes/No)
HECA score	Assessment sheet with inventory of high-energy hazards and binary assessment of Direct Controls for each hazard identified (1 = present control, 0 = absent control).	High-energy hazard identified: Fall from elevation Direct Control present for this hazard? (Yes/No)

Table 2
PJSB Quality Data Collection Instrument.

No.	Item Description
1	Everyone performing the job was present at the meeting
2	The discussion was held as close to the job as reasonably possible
3	Major work steps required to complete the job were identified and discussed
4	Necessary tools and equipment were identified and discussed
5	Hazards associated with the job were identified and discussed
6	Hazards posed by the environment or the surrounding work were identified and discussed
7	Controls for each identified hazard were identified and discussed
8	All life-threatening hazards and their controls were verbally differentiated and emphasized
9	Hazards and necessary controls were documented
10	All required permits were discussed before the work begins
11	Potential changes and were identified and discussed and a plan to address change was created
12	The importance of stopping work to address an unexpected change, or hazard was discussed
13	Emergency response plans were reviewed, including individual roles and responsibilities
14	Crew actively demonstrated their understanding of their work steps, hazards, and controls
15	All crew members participated in the discussion by identifying hazards and controls

Edison Electric Institute, 2024) and consisted of three videos of pre-job safety briefs and six images of different construction activities with their corresponding list of observations During each workshop activity, the participants individually rated each scenario using the standardized data collection tools, discussed the results of their collective ratings, and then identified differences in their ratings and opportunities for further alignment.

5.1.2. *Training assessment*

Before collecting field data, the assessors needed to validate their knowledge of the PJSB quality and HECA protocols against a standard. The validation involved obtaining both reasonable levels of consistency relative to their colleagues and precise scores relative to an answer key. To be consistent with their colleagues, participants needed to be within 0.5 of the absolute median deviation in the PJSB quality video exercise. Similarly, the assessors needed to be within two high-energy hazards apart from the HECA solution for their knowledge to be considered validated. These calibration thresholds had been validated previously (Bayona et al., 2025), and only the participants who met these thresholds were deemed as assessors and allowed to collect data in the field.

5.2. *Data collection*

Two main sources of data were required for this study: (1) field assessments on PJSB quality and HECA for various construction tasks within participant companies, and (2) injury records across the 31 construction companies. The PJSB quality and HECA data were collected over a period of three months in 2024. The injury records spanned 22 months, including 18 months preceding the PJSB quality and HECA data collection and 4 months following it.

To ensure the reliability of field assessments, assessor competency was calibrated and confirmed before data collection, while their recorded field assessments were subsequently validated during site assessments through pre-planned and independent assessments conducted by two researchers at each site. This data validation test assessed the variability in the field assessments provided by the trained assessors through the inter-rater reliability metric, which is a widely used tool to measure the agreement between two or more raters evaluating the same conditions (Koo & Li, 2016). Therefore, both PJSB quality and HECA data were validated before using them in further statistical tests.

5.2.1. *Measuring PJSB quality*

Once the assessors were trained and tested, they observed and assessed a series of pre-task safety meetings at multiple sites. During each observation, the assessors attended the pre-job briefings without intervening in the discussions or decisions made by the crew. Instead, they used the PJSB data collection instrument to check which of the 15 elements of a high-quality pre-job brief (see Table 2) were present or absent at the time of observation. The overall PJSB quality score for each safety brief was then computed by aggregating weighted scores based on the binary assessments (1 = present, 0 = absent) for each of the 15 scorecard elements. This computation yielded a continuous variable ranging from 0 to 100%, representing proportional values.

After collecting the PJSB quality data, its consistency was assessed using an inter-rater reliability metric. This was done to validate the reliability of the data gathered by the trained assessors. Cohen's kappa (κ) was used to quantify agreement above chance between assessors and researchers, across the 15 scorecard elements. Higher κ values indicated stronger levels of inter-rater reliability. For each company, an overall Cohen's kappa score was computed by analyzing the corresponding assessors' ratings.

5.2.2. *Measuring HECA*

In parallel with the pre-job brief assessments, the assessors evaluated different construction tasks on a regular basis over three months to collect HECA data. Each HECA corresponded to an observation of one or more workers within a crew doing a construction task. The HECA score was then computed as the ratio of the number of high-energy hazards with corresponding Direct Controls to all the high-energy hazards recorded at the time of observation. HECAs were done anytime crews were performing physical work, and the task observed did not need to be associated with the pre-job briefing evaluated. In some cases, the PJSB participants were not the same workers performing the tasks later observed. Therefore, unless the PJSB participants were the same workers who conducted the work later observed, the hazards discussed in the PJSB evaluation were different from those inspected in the HECAs. To preserve the traceability of the organization's injury statistics, the assessors were instructed to exclude any work performed by subcontractors. Therefore, only self-performed work was included in this study.

The HECA score data were computed on a continuous scale (0 – 100%). Therefore, HECA's reliability for multiple assessors in the same organization was measured through the Intraclass Correlation Coefficient (ICC). Similar to Cohen's approach, larger ICC values indicate strong agreement while lower (or negative) values indicate disagreement between raters, suggesting unreliable assessments. As suggested by Landis and Koch (1977) and Sun (2011), agreement values greater than 0.80 suggest near-perfect agreement, between 0.60 and 0.80 suggest substantial agreement, between 0.40 and 0.60 are regarded as moderate, between 0.20 and 0.40 as fair, and values below 0.20 indicate poor agreement.

The HECA inter-rater reliability was assessed at both the company and individual task levels. The combination of company and task-level analyses provided a more comprehensive view of the inter-rater reliability for heterogeneous organizations and diverse construction activities. To evaluate company-level reliability, the authors analyzed the assessments provided by all assessors from the same organization, including those by the two field researchers involved in the validation process. Then, for the task-level assessments, the authors compared each pair of researcher-assessor assessments to compute the overall Pearson correlation coefficient (ρ). Therefore, while the company-level estimates reflect the inter-rater reliability among all the employees in each participating company (i.e., calibration with their colleagues), the task-level estimates reflect the correlation of each researcher-assessor pair who observed individual tasks (i.e., validation against the researchers' assessments considered as ground truth).

5.2.3. *Obtaining injury records*

The injury records included the Occupational Safety and Health Organization's (OSHA) recordable injuries, the frequencies of first-aid injuries and SIFs, and the number of employee hours worked that a participant firm accrued 18 months preceding the data collection and four months post-data collection. This period established a representative link between the variables (Lingard et al., 2017).

These records corresponded to an archival data submission that was compiled by the research team and were needed to examine the statistical connections among the leading, monitoring, and lagging indicators. These data were retrieved from OSHA's Form 300A (OSHA. Occupational Safety and Health Administration, n.d.) and were entered by each participant company in a standardized format that contained each injury count and the corresponding worker hours monthly. While the number of first-aid injuries and OSHA recordable injuries were mutually exclusive, the number of SIFs corresponded to any other injury category. For example, an injury where a worker permanently lost a finger without being transferred or restricted from their regular duties could be classified as both a medical treatment case and a SIF (Bayona et al., 2023). Conversely, a non-SIF recordable injury, such as a cut in the hand requiring stitches, could only be classified into any of the medical treatment, job transfer or restriction, day-away-from-work case, nonfatal injury categories. Participating companies provided data on six lagging variables. These injury types corresponded to the number of:

(1) First-aid injuries
(2) Injuries that did not result in work transfer or lost time (i.e., medical treatment cases)
(3) Injuries resulting from job transfer or work restriction
(4) Injuries resulting in days away from work
(5) Fatal injuries
(6) SIFs

In summary, eight lagging (outputs), one leading (input), one condition (monitoring), and two firm demographic variables were selected as variables for this study (see Table 3). The demographic variables included company size (in terms of the number of hours worked per period) and the type of company. These variables were controlled by, given the potential varying injury effects due to company size (McVittie et al., 1997), and the different safety regulations between utility client companies and construction contractors.

5.3. *Data validation*

5.3.1. *Company-level assessment (PJSB quality and HECA)*

ICC estimates and their corresponding 95% confidence intervals were estimated using R statistical software version 20 (RStudio Team 2020) based on the average of k raters (usually ranging between 2 and 5), absolute agreement, and a two-way mixed-effects model. The Cohen's κ and ICC values for both PJSB quality and HECA for all participant companies evaluated are shown in Table 4. The results indicated moderately strong overall agreement for both PJSB quality and HECA. While there is room for improvement as indicated by some differences in their judgments, these agreement thresholds are considered acceptable in many scientific fields (Koo & Li, 2016; Alotaibi et al., 2024).

The inter-rater reliability results revealed varying levels of agreement between multiple assessors observing the same PJSB and construction tasks. For PJSB quality, substantial agreement was observed between the raters in 11 of 19 companies (58%), moderate agreement in 4 firms (21%), and fair agreement in 4 companies (21%). Similarly, for HECA, substantial agreement was seen in 10 out of 11 companies (91%), and fair agreement in 1 company (9%). No company reported poor reliability. Due to logistical challenges beyond the control of the research team, there were insufficient data to estimate stable Cohen's κ and ICC values for 2 and 10 companies, respectively.

Only companies with moderate to strong reliability were included in the analysis (i.e., κ or ICC greater than 0.40). This resulted in five companies being excluded from further statistical testing. Although this threshold may be low for clinical research (McHugh, 2012), this approach appears reasonable for similar reliability studies testing new protocols (Sim & Wright, 2005), especially when kappa values may have

Table 3
Variables examined.

Variable	Acronym	Definition	Variable type
Baseline pre-job safety brief quality	$\overline{PJSB}$	The average proportion of high-quality elements observed in a pre-job safety brief, based on a standard criterion.	Leading
Baseline HECA score	$\overline{HECA}$	The average proportion of high-energy hazards with a corresponding Direct Control observed for one crew performing a task.	Monitoring
Number of first-aid injuries	FA	Any injuries that do not require medical treatment and meet the set of criteria listed in OSHA's Form 300 (e.g., cleaning or flushing wounds on the skin surface).	Lagging
Number of injuries requiring medical treatment	MT	Any injury that requires medical treatment beyond first aid that does not meet the definition of job transfer or restricted, or days away from work.	Lagging
Number of job transfers or restricted case injuries	JT	A case where the employer or health care professional keeps or recommends keeping an employee from doing the routine functions of their job or from working a full workday that the employee would have been scheduled to work before the injury occurred.	Lagging
Number of days away from work injuries	DA	A case where an employee was away from work as a result of the recordable injury	Lagging
Number of fatal injuries	FT	An injury that resulted in death.	Lagging
Number of serious and fatal injuries	SIF	Any injury that resulted in death, near-death, or permanent injury (Bayona et al., 2023)	Lagging
Total recordable incident rate	TRIR	The rate at which a company experiences an OSHA-recordable incident, scaled per 200,000 work hours (Hallowell et al., 2021, p.28).	Lagging
Severity-based lagging indicator	SBLI	An adjusted injury rate that weights injuries by their relative severity and aggregates them into one rate (Hallowell and Oguz Erkal 2024, p.21)	Lagging
Company size	WH	Number of worker hours accumulated in the study period	Demographic
Company type	TYPE	Either a utility or a construction company	Demographic

been reduced by the high prevalence of positive agreements (i.e., how often people agreed on an element of a high-quality PJSB being present), measured through prevalence index metrics (see Table 4; Byrt et al., 1993).

5.3.2. *Task-level assessment (HECA)*

When evaluating HECA, the inter-assessor agreement at the company level was mirrored at the individual task level. A strong correlation was found between 55 task-rater pair assessments for 37 different tasks observed during the validation process, with a correlation coefficient of ρ (55) = 0.68, $p < 0.001$ (95% CI = 0.51——0.80). In line with the company-level assessments, these results suggest moderately strong reliability between raters when assessing the same task.

Table 4
Inter-rater reliability estimates and their confidence intervals.

Company	No. observers	Cohen's κ	PI κ	95% CI	ICC	95% CI
1	2	0.87*	0.00	0[.62, 1.0]	–	–
2	2	0.53*	0.40	0[.06, 0.99]	1.00	[1.0, 1.0]
3	4	0.43*	0.27	0[.29, 0.74]	–	–
4	2	0.70*	0.33	0[.32, 1.0]	0.31†	[-9.1, 0.98]
5	4	–		–	0.71*	[3.7, 0.99]
6	2	0.47	0.53	0[.02, 0.93]	0.75	[-9.1, 1.0]
7	2	0.21†	0.40	[-0.31, 0.73]	0.96	[-0.54, 1.0]
8	2	0.86*	0.46	0[.59, 1.0]	0.95	[-0.61, 1.0]
9	2	0.72*	0.20	0[.36, 1.0]	–	–
10	2	0.21†	0.06	[-0.26, 0.69]	–	–
11	4	0.32†	0.73	0[.00, 0.76]	–	–
12	2	1.00*	0.33	[1.0, 1.0]	0.96*	0[.63, 0.99]
13	3	0.90*	0.26	0[.86, 1.0]	–	–
14	3	0.65*	0.53	0[.47, 1.0]	0.87	[-0.17, 0.99]
15	5	0.36†	0.26	0[.30, 0.87]	–	–
16	2	0.81*	0.53	0[.47, 1.0]	0.98*	0[.57, 1.0]
17	3	0.91*	0.13	0[.86, 1.0]	0.61	[-0.26, 0.99]
18	4	0.60*	0.40	0[.24, 0.86]	–	–
19	2	0.81*	0.53	0[.47, 1.0]	–	–
20	2	–	–	–	0.94*	0[.69, 0.99]
21	2	0.73*	0.07	0[.39, 1.0]	–	–

Note. CI = Confidence interval, PI = Prevalence index. * Denotes statistical significance at the 5% alpha level. † Denotes the companies that were excluded from further analysis.

5.3.3. *Data cleaning and pre-processing*

Once the raw data were collected, it was scrutinized and pre-processed into a consistent and usable format for further processing. First, any out-of-scope data, such as information about low-energy hazards, were removed from the field data collection forms. Then, partner companies were asked to reconcile any missing data, structural errors, and other inconsistencies in the injury records (e.g., some companies reported the number of days rather than the number of cases of days-away-from-work injuries). After addressing the data quality issues, a data diagnostic including leverage and Cook's distance values was conducted to identify and remove outliers. As a result, data from three companies were excluded from the analysis.

Once the data were cleaned, it was manually entered into a consistent format. Baseline measures for both PJSB quality and HECA were then computed by calculating the sample average of the field assessments at each company and assessing the stability of the ratios between the sample standard deviation to the sample mean (i.e., coefficients of variation or CV). These measures provided a reference measure of the current state of safety of the participant companies. After that, the injury records were converted into broad injury rates by normalizing them to the number of worker hours reported by each organization. Two broad injury rates were used as lagging indicators in this study: TRIR and SBLI.

Separate injury rates were calculated for each “i” company, considering both pre- and post-field data collection periods (18 and 4 months, respectively). These injury rates were calculated using the formulas described below:

$$TRIR_i = \frac{(MT_i + JT_i + DA_i + FT_i)*200,000}{WH_i} \tag{1}$$

$$SBLI_i = \frac{(100*FA_i + 500*MT_i + 750*JT_i + 1500*DA_i)*200}{WH_i} \tag{2}$$

5.3.4. *Regression models and statistical testing*

Descriptive statistics and goodness of fit tests helped determine statistical models that best fit the data. Equal variance (homoscedasticity) and normality assumption tests showed that linear regression models were suited to test the null hypotheses of no relationship among PJSB quality, HECA, and historical injury rates. Additionally, overdispersion tests and relative frequency histograms showed that most post-data collection injuries had excessive zeros (see Fig. 3). Therefore, linear regression models were used to test links among *past* injuries, PJSB quality, and HECA, whereas generalized linear models (GLM) tested the relationship between both PJSB quality and HECA, with *future* injuries. A Bayesian Information Criterion (BIC) and Likelihood Ratio goodness-of-fit tests revealed that a Poisson count model outperformed more complex zero-inflated models that accounted for excessive zeros ($p < 0.02$), which are common in injury data (e.g., Versteeg et al., 2019). The regression models used to test the hypotheses are shown in Table 5.

As indicated in Table 5, Model 1 tested the mutual link between baseline PJSB quality and HECA scores, and Models 2 and 3 tested the relationship among historical injuries, PJSB quality, and HECA. To account for variability in company size and differences in safety regulations between utility clients and contractors, Model 1 included the log-transformed hours worked, while Models 2 and 3 included the type of company as a covariate. Analysis of Variance (ANOVA) and BIC tests verified the relative contributions of these additional parameters to the models. Lastly, Model 6 predicted the number of injuries of a specific type (i.e., FA, MT, JT, etc.), adjusting for the company size.

Table 5
Regression models tested.

Model	Equation
1	$HECA_i = \beta_0 + \beta_1 \overline{PJSB}_i + \beta_2 \log(WH_i) + \varepsilon_i$
2	$PJSBquality_i = \beta_0 + \beta_1 TRIR/SBLI_i + \beta_2 TYPE_i + \varepsilon_i$
3	$HECA_i = \beta_0 + \beta_1 TRIR/SBLI_i + \beta_2 TYPE_i + \varepsilon_i$
4	$TRIR_i = \beta_0 + \beta_1 \overline{PJSB}/\overline{HECA}_i + \varepsilon_i$
5	$SBLI_i = \beta_0 + \beta_1 \overline{PJSB}/\overline{HECA}_i + \varepsilon_i$
6	$log\left(\frac{E[InjurySeverity]_{hi}}{WH_i}\right) = \beta_0 + \beta_1 X_1 + \log(WH_I)$

Note. Within the models, *h* represents an *Injury Severity* type (either FA, MT, JT, DA, FT, or SIF), *i* represents the participant company, Injury Severity is the expected value of the type of injury for company *i*, and $\beta_1 X_1$ is the term for the predictor variable (PJSB or HECA) with their estimated coefficients (β_1).

6. Results

6.1. *Descriptive statistics*

The final sample consisted of 999 PJSB quality and HECA field evaluations drawn from 278 companies: 11 in heavy civil construction (41%), 10 in electrical, gas, and water utilities (37%), and 6 in the commercial building sector (22%) (see Table 6). The distribution of pre-job quality and HECA scores for these companies is shown in Fig. 4. Collectively, these firms accounted for 361 million worker hours in the 22-month study period, where injuries were reported. As indicated in Table 6, the final sample included a total of 4,992 first-aid injuries, 1,092 medical treatment, 715 job transfer or restriction, 982 lost-time, 4 fatal injuries, and 106 SIFs. On average, participant companies reported a TRIR and SBLI of two injuries per 100 full-time workers.

6.1.1. *Baseline measures of PJSB quality and HECA*

Baseline PJSB quality ranged from 0.25 to 0.89, whereas baseline HECA scores ranged from 0.17 to 0.92 (see Table 6). Baseline PJSB quality had a mean and standard deviation of 0.62 and 0.15, while baseline HECA scores had a mean and standard deviation of 0.53 and 0.21. Except for two companies, all participating firms provided a data sample that matched or exceeded expectations. This limitation did not

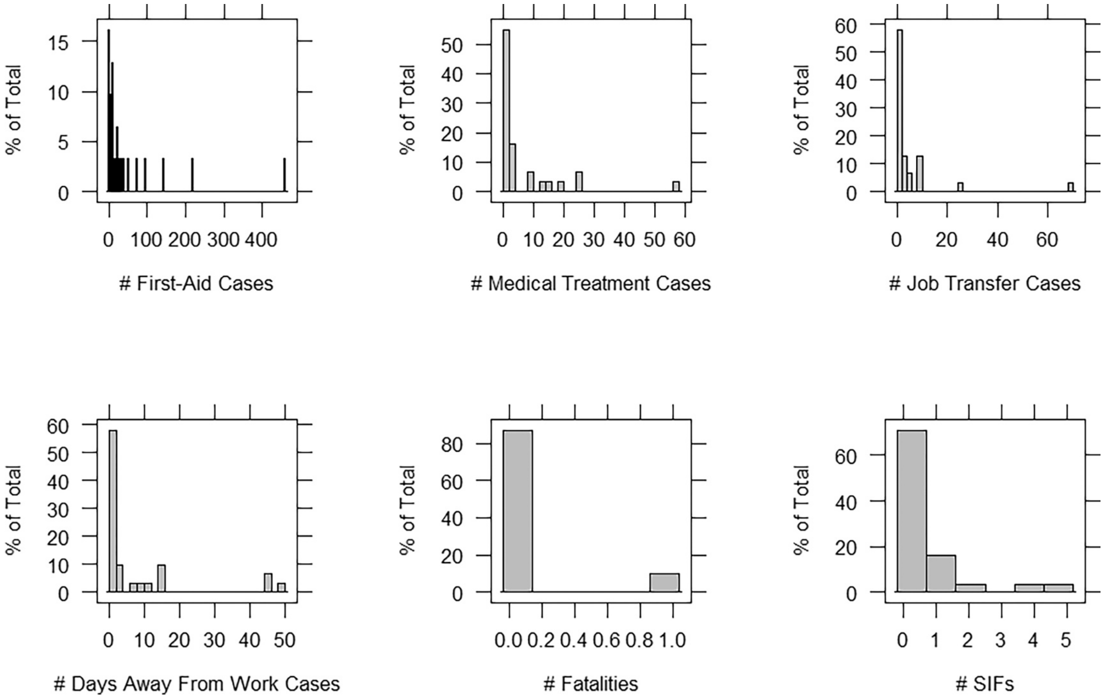

Fig. 3. Relative frequency histograms for post-data collection injuries.

Table 6
Descriptive statistics for study variables.

Variable	N	Mean	SD	Minimum	Maximum	Total
PJSB quality (count)	19	21.83	7.16	12	41	439
HECA (count)	27	21	6.25	11	37	560
$\overline{PJSB}$	19	0.62	0.16	0.25	0.89	–
$\overline{HECA}$	23	0.54	0.21	0.17	0.92	–
WH (in thousands)	31	11,648	16,693	42	74,916	361,109
FA (count)	31	161.03	275.15	0	1251	4992
MT (count)	31	35.23	77.52	0	415	1092
JT (count)	31	23.06	41.25	0	187	715
DA (count)	31	31.68	60.60	0	248	982
FT (count)	31	0.13	0.34	0	1	4
SIFs (count)	31	3.42	6.69	0	29	106
TRIR[a]	31	2.00	2.05	0	9.55	–
SBLI[a]	31	2.08	1.76	0	6.95	–

[a] These rates were normalized using 200,000 worker hours for the entire study period (22 months).

invalidate the baseline measures of PJSB quality and HECA, as the work performed by the limited-sample companies was relatively stable and consistent (i.e., earthwork, micro-tunneling, and site cleanup).

Coefficients of variation (CV), computed as the ratio of the sample standard deviation to the mean, ranged from 18% to 30% for PJSB quality and from 18% to 95% for HECA, indicating moderate to high variability in the field observations. 95%-confidence intervals for the mean ranged between 10% and 50% of the mean across all measurements, suggesting a reasonable level of precision for PJSB quality but a broader range for the HECA scores. These results support the use of mean statistics as a more reliable measure of baseline PJSB quality than HECA scores. While variability was present in the data, it remained within acceptable thresholds commonly observed in safety research (Alotaibi et al., 2024).

6.1.2. Relationship between PJSB quality and the HECA score

Table 7 shows the regression results for Model 1 (see list of regression models in Table 5), which examined the relationship between PJSB quality and HECA scores across participant firms. Within this model, baseline PJSB quality values were significantly associated with baseline HECA (CI = 0.02, 0.96; p = 0.04), controlling for the number of employee hours. Practically, each percentage point increase in PJSB quality could yield, on average, a 0.5% increase in HECA score (i.e., proportion of high-energy hazards with a Direct Control). This evidence supports the hypothesis H1, suggesting that organizations with higher-quality pre-job safety briefs tend to have more Direct Controls in place.

6.1.3. Relationships among historical injury rates, PJSB quality, and HECA

The results of Models 2 and 3 are also presented in Table 7, where TRIR and SBLI were used to explain variations in both HECA and PJSB quality. Statistically significant associations were observed between PJSB quality and TRIR (p = 0.03), and between PJSB quality and SBLI (p = 0.03), after controlling for the type of company (utility client vs.

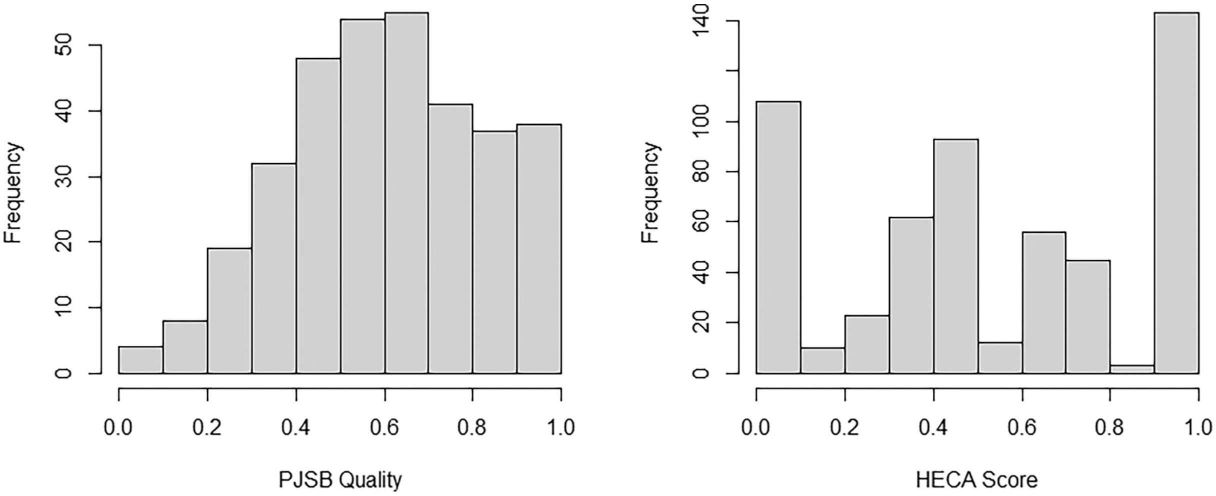

Fig. 4. Distribution of pre-job safety brief and HECA field evaluations.

Table 7
Linear regressions of links between PJSB quality, the HECA score, and injury rates.

Predictor	Target variable	Parameter Estimate	SE	t	p-value	95% CI	Adj R^2
Model 1							
$\overline{PJSB}$	$\overline{HECA}$	0.49	0.22	2.22	**0.04**	0[.020, 0.969]	0.56
Model 2							
TRIR	$\overline{PJSB}$	−3.60	1.57	−2.28	**0.03**	[-6.95, −0.24]	0.23
SBLI	$\overline{PJSB}$	−4.60	1.98	−2.32	**0.03**	[-8.82, −0.37]	0.24
Model 3							
TRIR	$\overline{HECA}$	−3.78	1.52	−2.47	**0.02**	[-6.93, −0.62]	0.39
SBLI	$\overline{HECA}$	−3.84	1.95	−1.97	0.06	[-7.87, 0.183]	0.34
Model 4							
$\overline{PJSB}$	TRIR	−0.04	0.05	−0.69	0.49	[-0.158, 0.083]	−0.03
$\overline{HECA}$	TRIR	−0.01	0.03	−0.50	0.62	[-0.081, 0.049]	−0.03
Model 5							
$\overline{PJSB}$	SBLI	−0.05	0.08	−0.65	0.52	[-0.227, 0.120]	−0.03
$\overline{HECA}$	SBLI	−0.02	0.04	−0.63	0.53	[-0.119, 0.063]	−0.02

Note. CI = confidence interval, Adj R^2 = adjusted R-squared.

contractor). A significant association was also seen between TRIR and the HECA score (CI = -6.93, −0.62; $p = 0.02$). These results provide support for the hypotheses H2 that organizations with fewer past injuries tend to have better PJSB quality and HECA.

These results highlight the role of TRIR and SBLI in explaining the variation in PJSB quality and HECA. Each additional OSHA recordable per 200,000 worker hours was associated with an approximate 4%-reduction in PJSB quality and HECA values. The negative link between SBLI and pre-job briefs suggests that organizations experiencing a higher frequency of severe injuries resulting in job transfer, restrictions, or lost time tend to have less-focused pre-job safety meetings. This finding reinforces the idea that better pre-task planning is associated with a reduced incidence of lost-time or work-restricted injuries.

6.1.4. Relationships among PJSB quality, HECA, and long-term injuries

There was a negative relationship between both baseline PJSB quality and HECA, and TRIR/SBLI, suggesting that companies with better pre-job briefs and more Direct Controls experience fewer long-term rates of recordable injuries. However, as shown in Table 7, none of these linear relationships were significant at the 5% alpha level. Although this result does not support hypotheses H4, it only constitutes partial evidence, as the models had a poor fit, as seen by the low and negative adjusted R^2 values.

6.1.5. Differences in PJSB quality and HECA between clients and contractors

The significant association between historical injury rates (TRIR and SBLI) and baseline PJSB quality were only found when controlling for the type of company (client or contractor), suggesting that the type of company was confounding the true association between PJSB quality and TRIR and between PJSB quality and SBLI. As indicated in Fig. 5, on average, clients reported higher PJSB quality and HECA scores than contractors ($\Delta = 22.5\%$), suggesting differences in the effectiveness of safety management practices across different types of firms. Therefore, there is evidence to support H3, which posited that certain types of companies have better PJSB quality and more Direct Controls than others. Specifically, the results showed that utilities have, on average, higher PJSB quality and higher HECA scores than contractors. Interestingly, companies with near-zero PJSB scores (i.e., no formal pre-job safety briefing) consistently lacked Direct Controls in place. On the flip side, companies with near-perfect PJSB quality did not achieve perfect baseline HECA scores, suggesting that other organizational factors may influence the implementation of Direct Controls among high-performing companies.

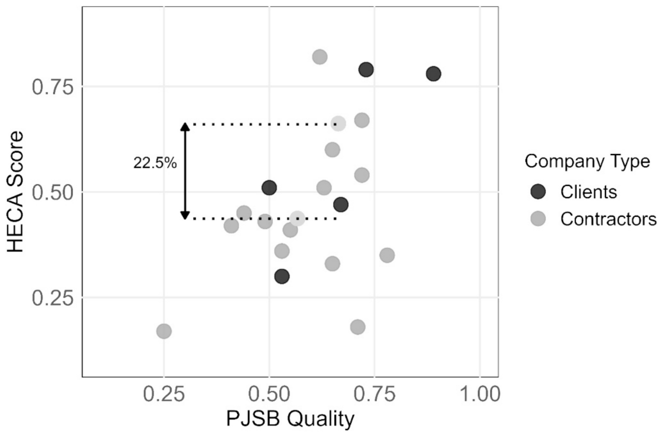

Fig. 5. ***Differences in PJSB quality and HECA scores between utility clients and contractors.*** *Note.* The light gray dots show the overall average baseline PJSB quality and HECA scores across the utilities and contractor companies.

6.2. GLM results

6.2.1. Non-linear relationships among PJSB quality, HECA, and long-term injury rates

The GLM results for Model 6 are shown in Table 8, where baseline PJSB quality and HECA scores predicted the expected counts of specific injury types (i.e., first aids, OSHA recordables, and SIFs) for four months' post-field data collection. This was done to obtain a more stable relationship between the short-term measures of safety and the long-term injury rates. Within these models, the relationship between baseline HECA scores and every injury type was statistically significant at the 5% alpha level. This non-linear regression approach more accurately represented the distribution of the count (injury) data for more severe injuries, as seen by the relatively low root mean square errors for fatal injuries and SIFs (see Table 8). This suggests that companies with higher baseline HECA scores may expect a reduced number of long-term SIFs than firms with higher exposure to high-energy hazards. This was also true for baseline PJSB quality and first aids, and for baseline PJSB quality and job transfer or work restriction cases. Despite this significance, the high Root Mean Square Error (RMSE) values for less severe injuries suggested a high degree of error in the injury predictions. There was also a significant relationship between PJSB quality and SIFs, but in the opposite direction than expected. Companies with higher baseline PJSB quality showed an increase in long-term SIFs. One possible explanation is that organizations with more formal pre-job briefings may have greater transparency and consistency in reporting SIFs, which leads to a higher number of reported cases.

As indicated in Table 8, HECA is a strong predictor of future SIF risk. Every percentage increase in a company's baseline HECA score reduces the expected number of SIFs by 3%, holding work hours constant (CI = -0.05, − 0.001; $p = 0.03$). Therefore, these results support H4, particularly for SIFs, suggesting that companies with more Direct Controls may expect fewer long-term SIFs.

7. Discussion

This paper explored the statistical relationships among PJSB quality (a leading indicator), HECA (a monitoring indicator), and injury rates (a lagging indicator). As indicated in Fig. 6, findings revealed: (1) a positive connection between the two short-term metrics (PJSB quality and HECA), (2) a negative correlation between historical TRIR/SBLI and PJSB quality, and between TRIR and HECA, (3) and a negative association between the HECA score and the expected number of long-term SIFs. These findings are correlational and do not establish causal links. Specifically, for every percentage point increase in pre-job safety brief quality, there was a 0.5% increase in HECA, and a 3% reduction in the expected number of SIFs, holding the number of employee worker hours constant. Typical baseline PJSB quality scores ranged from 0.25 to 0.89, while baseline HECA scores ranged from 0.17 to 0.92, with overall company averages of 0.62 and 0.53, respectively.

The following takeaways stem from our empirical findings depicted in Fig. 6.

Finding 1: There is a connection between PJSB quality and HECA (two short-term metrics).

Findings revealed a positive relationship between SIF-focused pre-job safety meetings and the proportion of Direct Controls for high-energy hazards across different firms. As anticipated, companies that conduct more structured and engaged pre-job safety briefs tend to have more Direct Controls in place. Specifically, for every percent increase in pre-job quality, there was a 0.5% increase in HECA, holding worker hours constant. This proactive approach to safety reflects a strong culture, where structured pre-job meetings and improved controls are part of a company's strategies to prevent SIFs.

The observed link between the two unique precursors of SIFs contributes to existing literature by demonstrating the merits of participatory pre-task planning activities on increasing the levels of Direct

Table 8
Generalized linear model results.

Predictor	Dependent variable	Parameter estimate	SE	z-value	Wald 95% CI	*p*-value	RMSE
Model 6							
$\overline{\textit{PJSB}}$	FA	−0.02	0.00	−8.89	[-0.03, −0.02]	**$2\text{x}10^{\hat{}-6}$**	58.97
	MT	0.01	0.00	1.14	[-0.07, 0.02]	0.25	17.35
	JT	−0.03	0.00	−4.89	[-0.05, −0.02]	**$1\text{x}10^{\hat{}-6}$**	16.44
	DA	0.01	0.00	1.92	[-3.49, 0.03]	0.054	21.19
	FT	−0.04	0.06	−0.71	[-0.13, 0.10]	0.47	0.23
	SIF	0.07	0.03	2.38	[0.01, 0.14]	**0.01**	1.40
$\overline{\textit{HECA}}$	FA	−0.02	0.00	−11.9	[-0.02, −0.01]	**$2\text{x}10^{\hat{}-16}$**	41.91
	MT	−0.01	0.00	−3.58	[-0.02, −0.00]	**$3\text{x}10^{\hat{}-4}$**	14.08
	JT	0.04	0.00	6.11	[0.02, 0.05]	**$9\text{x}10^{\hat{}-10}$**	12.73
	DA	−0.01	0.00	−4.85	[-0.02, −0.01]	**$1\text{x}10^{\hat{}-6}$**	17.50
	FT	−0.08	0.03	−2.68	[-0.15, −0.02]	**$7\text{x}10^{\hat{}-3}$**	0.28
	SIF	−0.03	0.01	−2.27	[-0.06, −0.00]	**0.02**	1.31

Note. SE = standard error, RMSE = Root Mean Square Error. FA = number of first-aid cases, MT = number of medical treatment cases, JT = number of job transfer or restricted work cases, DA = number of days-away-from-work cases, FT = number of fatalities, SIF = number of serious and fatal injuries.

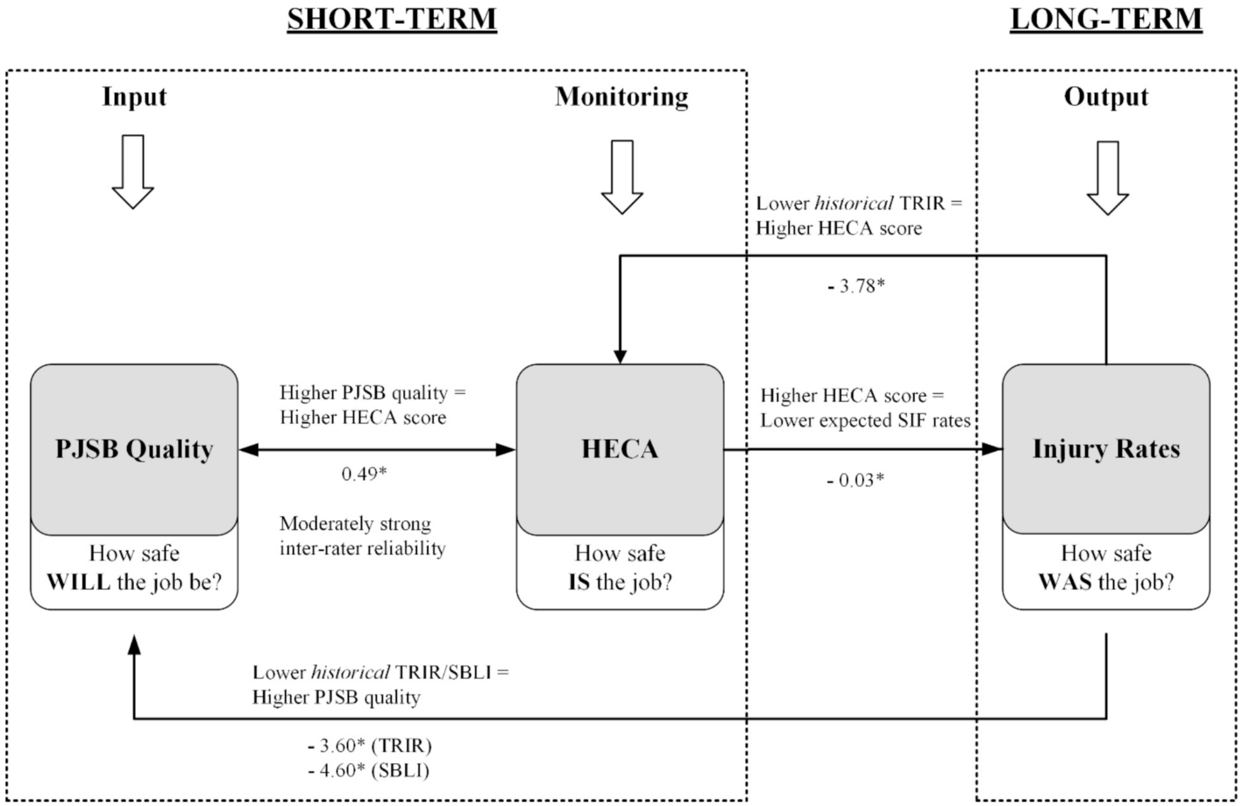

Fig. 6. Summary of results. *Note.* Coefficients presented are linear regression coefficients. = Baseline pre-job safety brief quality for each company, HECA = Baseline HECA score for each company, TRIR = Total Recordable Incident Rate, SBLI = Severity-Based Lagging Indicator, SIF = Number of serious injuries and fatalities. **p < 0.05.

Controls, which are critical for SIF prevention. Prior studies indicate that participatory safety activities: (1) increase injury risk awareness (Swartz, 2002; Roughton & Crutchfield, 2015), (2) improve the understanding of safe work procedures and safety concerns (Rausand & Haugen, 2020), (3) foster a sense of ownership and positive safety attitudes (Roughton, 2002), and (4) promote more effective hazard controls (Zheng et al., 2017; Albrechtsen et al., 2019). This research builds on that foundation by demonstrating that higher-quality engagement in PJSB correlates with increased crew awareness and improved control measures for hazardous energy sources directly related to SIF conditions.

Improved decision-making and crew situational awareness may partially explain the observed association between the two short-term safety metrics as well (e.g., Roughton & Crutchfield, 2015; Zheng et al., 2017; Albrechtsen et al., 2019). For example, the PJSB data collection tool prioritizes the identification of hazardous energy sources that lack adequate safeguards, prompting teams to discuss gaps in safe work plans (Albrechtsen et al., 2019). Additionally, incorporating worker input creates opportunities to collectively surface safety concerns and co-develop alternative safeguards against life-threatening hazards (Memarian et al., 2023). Field supervisors may incorporate feedback from pre-job discussions into other safety planning protocols such as Job Hazard Analyses and Field-Level Risk Assessments to achieve improved safety performance. Crew leaders and safety supervisors may also review energy transfer mechanisms for specific tasks to recommend countermeasures such as new controls, improved barriers, or specialized training (Rausand & Haugen, 2020).

Finding 2: Organizations with lower injury rates tend to have better short-term safety performance (i.e., higher PJSB quality and HECA).

There is an important association between historical injury rates and short-term safety performance, but in the opposite direction that some would expect. Findings showed that companies may respond to lower historical injury rates by improving their pre-job safety brief activities. Specifically, for every additional OSHA recordable injury, companies experienced a 3.6% reduction in PJSB quality and a 3.7% reduction in HECA. This result challenges previous findings by Lingard et al. (2017), who found that companies throttle some safety activities like toolbox talks in response to poor injury performance. Nonetheless, this finding corroborates the complexity in the statistical links between safety activities and injury rates found within the literature (e.g., Versteeg et al., 2019).

The lack of a significant association between historical SBLI rates and HECA may reveal some gaps in how a stronger safety culture translates to better safety controls during active work. While disconcerting, this result is consistent with previous research by Kang et al. (2017), who found that 88% of construction fall accident victims lacked proper fall protection systems.

Finding 3: The HECA score is a strong predictor of future fatality risk.

Findings demonstrate that HECA is a strong predictor of high-severity injuries, including SIFs (See Fig. 7). The results showed that the likelihood of serious and fatal injuries significantly decreased with improved levels of Direct Controls. For every percent increase in HECA, companies expected a 3% reduction in the expected number of SIFs, holding worker hours constant. This result is consistent with previous literature, as energy-focused controls have been highly effective in reducing the risk of fatal crash injuries. Since their inception, airbags and child restraint systems have saved over 60,000 lives altogether (Kahane, 2015; Ehsani et al., 2023). These types of controls have also been effective at significantly reducing the severity of injuries caused by roadway crashes, with some barriers like median cable barriers being more effective at reducing injury risk than others such as guardrails (Zou et al., 2014). These findings corroborate the study results and support the need for more effective hazardous energy barriers in construction.

The following example illustrates how modest improvements in an organization's baseline HECA score (i.e., its stable average) could lead to meaningful reductions in serious and fatal injury risk. As seen in Fig. 7, holding exposure steady at 2 million employee hours worked (i.e., the median number of worker hours in the study sample, representing roughly 1,000 full-time employees [FTE] in a year):

- A firm with a HECA score of 0 would be expected to experience around 2.5 SIFs per 1,000 FTE
- At HECA 0.50, that number drops to about 0.50
- At HECA 0.90, the model predicts less than 0.15 SIFs.

The results showed that PJSB quality is a relatively stable proxy for HECA and long-term injury rates. Therefore, companies may continuously track both PJSB quality and HECA to identify high-energy hazards and assess Direct Controls. Although the necessary frequency of these assessments may vary based on work complexity and company size, it is recommended that organizations obtain baseline averages that at least match the sample size from this study, in a short-term period of no more than three months. This implies obtaining at least 15 assessments of PJSB quality and 15 HECAs gathered randomly for different tasks and crews. Furthermore, while a minimum unacceptable threshold for HECA is still uncertain, organizations are encouraged to provide coaching to crews and address gaps in existing controls if HECA scores are less than 30%, at which point a SIF becomes a likely event (See Fig. 7).

8. Study limitations

While the statistical methods provided robust insights into the relationships between the short-term metrics and long-term injury rates, several limitations warrant consideration when interpreting the results. First, the scope of the data imposed certain restrictions on the generalizability of the findings, particularly regarding a large, convenient sample of volunteering companies, insufficient data to run a within-company or utility versus construction comparative analysis, and the different reporting patterns for first-aid injuries. Although the under-reporting of minor injuries has been widely discussed in the literature (e.

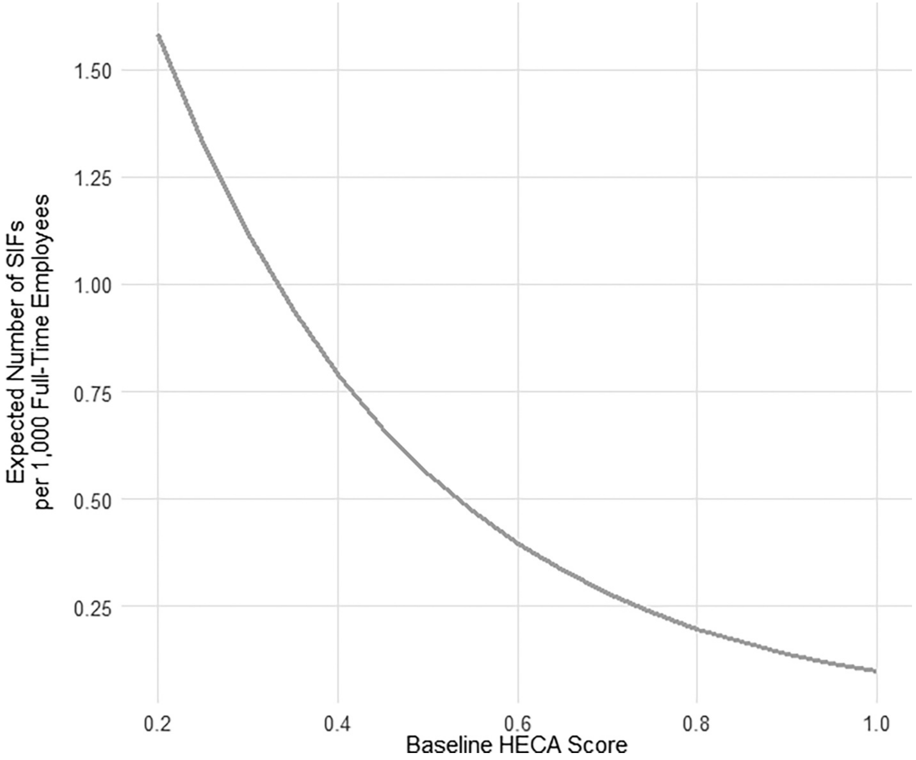

Fig. 7. Predicted number of fatalities per 1000 full-time employees within a company.

g., Versteeg et al., 2019), the significant effects encountered in more severe injuries alleviate this issue. Furthermore, although the use of GLM regression enabled nuanced predictive modeling, it did not account for the temporal variability of the injury data or unobserved confounders that may have explained the relationships found.

Several methodological limitations stem from the nature of data collection processes involving industry professionals. For instance, Hawthorne effects may have influenced safety-related practices during site evaluations. However, Hawthorne effects were likely constant, as most of the observers were company employees who were familiar with the employees observed. Additionally, reliance on subjective PJSB quality and HECA assessments by industry professionals introduced assessment variability that reflected individual experience, interpretive frameworks, and organizational norms rather than purely objective hazard identification. This in turn, reduced the inter-reliability of these metrics. Lastly, this study did not establish causal relationships. It is unclear whether fewer injuries originate from more Direct Controls and more engaged PJSBs or whether firms have more Direct Controls and better PJSBs because of lower injury rates. The directionality of the links established should be further examined through longitudinal studies involving a large sample size.

9. Conclusions and recommendations

The prevention of serious and fatal injuries has become a priority in the construction industry. Concerned by the persistent plateau in fatal injury rates, organizations are increasingly seeking alternative evidence-based strategies to reduce exposure to SIF risks. Emerging safety science has shown that planning for and controlling high-energy hazards is crucial to reducing the risks of fatal injury. Therefore, this research explored and validated a holistic and balanced measurement framework where novel short- and long-term safety metrics were described in three temporal themes: as inputs, as monitoring activities, and as outputs of past performance. Thus, this study examined two differentiators of SIFs, namely PJSB quality and HECA.

Findings revealed that companies with lower injury rates had more engaged pre-job safety briefs and higher HECA scores than low-performing firms. The authors also found a relevant connection between PJSB quality and HECA, and a reduced number of future injuries, including SIFs, in high-performing companies. These findings have relevant methodological and practical implications. Methodologically, this research protocol demonstrates the ability to generate large volumes of rigorous, mostly reliable safety data through industry-academic collaboration and standardized training protocols. Practically, the results may confirm the veracity of the theoretical connections between quality-based leading, monitoring, and lagging indicators in the context of SIF prevention. These connections further align with modern definitions of safety that coexist under a logical multi-phase approach that views safety as inputs to reach desired goals, monitoring of safe conditions, and examinations of past outcomes that reflect the predicted levels. Overall, the research reveals a general correspondence among pre-task planning focused on SIF risks, strong levels of Direct Controls during ongoing work, and better long-term safety performance, as seen by the correspondence between improved short-term performance and lower rates of long-term injuries.

This paper provides early empirical evidence that companies can measure safety differently to obtain a more comprehensive story about their safety. They can measure safety more stably and consistently by measuring the presence of controls and predicting long-term outcomes. This is particularly relevant for companies transitioning from tracking lagging indicators only to a more balanced suite of leading-monitoring-lagging metrics that collectively address the limitations of traditional safety performance assessment.

Future studies can further examine the causal link between the unique precursors of SIFs and long-term safety performance. Understanding this link can further elucidate why organizations are experiencing a larger number of serious injuries, and where managers could intervene to improve critical conditions that lead to SIFs. Additionally, future research can explore whether utility clients have better short-term safety performance due to the high-risk nature of their business or due to different safety regulations. A holistic approach to safety measurement and a deeper understanding of the mechanisms leading to SIFs may ultimately contribute to reducing the disconcerting plateau of fatal injuries in the industry.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to improve the clarity and readability of a few sections of the paper. The content was originally written by the authors and ChatGPT was used to check for grammar, format, and style. ChatGPT was used to clarify writing, but not to create content. After using this tool, the authors carefully reviewed and edited the content as needed and take full responsibility for the content of the published article.

CRediT authorship contribution statement

Arnaldo Bayona: Writing – original draft, Investigation, Methodology, Data curation, Visualization. **Matthew R. Hallowell:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Siddharth Bhandari:** Writing – review & editing, Supervision. **Yaqoob Raheemy:** Writing – review & editing, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was made possible through the support of the Construction Safety Research Alliance (CSRA). The authors thank the CSRA industry leaders and all the safety professionals who participated in this study for providing data, advice, and inspiration. This work would not have been possible without their contributions.

References

Albrechtsen, E., Solberg, I., & Svensli, E. (2019). The application and benefits of job safety analysis. *Safety science, 113*, 425–437.

Alexander, D., Hallowell, M., & Gambatese, J. (2017). Precursors of construction fatalities. I: Iterative experiment to test the predictive validity of human judgment. *Journal of construction engineering and management, 143*(7), Article 04017023.

Alotaibi, A., Gambatese, J., & Nnaji, C. (2024). Developing a novel energy-based approach for measuring mental workload. *Heliyon*. https://doi.org/10.1016/j.heliyon.2024.e24828

Alruqi, W. M., & Hallowell, M. R. (2019). Critical success factors for construction safety: Review and meta-analysis of safety leading indicators. *Journal of construction engineering and management, 145*(3), Article 04019005. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001626

Anderson, M., and Denkl, M. (2010, April). The Heinrich accident triangle—too simplistic a model for HSE management in the 21st century?. In *SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability?* (pp. SPE-126661). SPE.

Balderson, D. (2016). Safety defined: A means to provide a safe work environment. *Professional Safety, 61*(05), 63–68.

Ball, D. R., & Frerk, C. (2015). A new view of safety: Safety 2. *British Journal of Anaesthesia, 115*(5), 645–647.

Bayona, A., Bhandari, S., Hallowell, M., Sherratt, F., Bailey, J. M., & Upton, J. (2023). What is a serious injury? a model for defining serious injuries & fatalities. *Professional Safety, 68*(09), 22–30.

Bayona, A., Hallowell, M. R., & Bhandari, S. (2024). The things that hurt people are not the same as the things that kill people: key differences in the proximal causes of low- and high-severity construction injuries. *Journal of Construction Engineering and Management, 150*(8), Article 04024089.

Bayona, A., Hallowell, M. R., Bhandari, S., Moyen, N., & Lien, A. (2025). Impact of energy-based safety training on quality of prejob safety meetings and control of hazardous energy in construction: multiple baseline experiment. *Journal of Construction Engineering and Management, 151*(7), Article 04025086.

Bayona Malo, A.A. (2025). *Evidence-Based Strategies to Define, Differentiate, and Prevent Serious Injury and Fatality Exposure in Construction* (Doctoral dissertation, University of Colorado at Boulder).

Bayramova, A., Edwards, D. J., Roberts, C., & Rillie, I. (2023). Constructs of leading indicators: a synthesis of safety literature. *Journal of safety research, 85*, 469–484.

Bellamy, L. J. (2015). Exploring the relationship between major hazard, fatal and nonfatal accidents through outcomes and causes. *Safety Science, 71*, 93–103.

BLS (Bureau of Labor Statistics). (2025). *Fatal work injuries fell in 2023*. U.S. Bureau of Labor Statistics. Retrieved May 16, 2025, from https://www.bls.gov/opub/ted/2025/fatal-work-injuries-fell-in-2023.htm?utm_source=chatgpt.com.

Busch, C., Usrey, C., Loud, J., Goodell, N., & Carrillo, R. A. (2021). Serious injuries & fatalities: Why are they constant while injury rates decrease? *Professional Safety, 66*(01), 26–31.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of clinical epidemiology, 46*(5), 423–429.

Conklin, T. (2019). *The 5 Principles of Human Performance: A contemporary update of the building blocks of Human Performance for the new view of safety*. Pre-Accident Investigation Media.

Cooper, M. D. (2019). The efficacy of industrial safety science constructs for addressing serious injuries fatalities (SIFs). *Saf. Sc., 120*, 164–178. https://doi.org/10.1016/j.ssci.2019.06.038

CSRA (Construction Safety Research Alliance). (n.d.). *Quality of Safety Leading Indicators*. CSRA. https://www.csra.colorado.edu/qsli.

Dekker, S. (2014) Safety differently: Human factors for a new era. CRC Press.

EEI (Edison Electric Institute). (2024, February 16). *Severity-Based-Lagging-Indicator – Making the Best of Our Injury Data.* The Power to Prevent Serious Injuries & Fatalities. https://www.eei.org/en/issues-and-policy/power-to-prevent-sif.

EEI (Edison Electric Institute). (2025, July 3). *Issues & Policy. The Power to Prevent Serious Injuries & Fatalities.* Edison Electric Institute. https://www.eei.org/en/issues-and-policy/power-to-prevent-sif.

Ehsani, J. P., Michael, J. P., & MacKenzie, E. J. (2023). The future of road safety: challenges and opportunities. *The Milbank Quarterly, 101*(Suppl 1), 613.

Golabchi, H., Abellanosa, A. D., Lefsrud, L., Pereira, E., & Mohamed, Y. (2024). A comprehensive systematic review of safety leading indicators in construction. *Safety Science, 172*, Article 106433.

Hallowell, M. R., & Oguz Erkal, E. D. (2024). Severity-based lagging indicator: an alternative measure of safety performance. *Professional Safety, 69*(04), 20–27.

Hallowell, M. R., & Spencer, C. (2024). Safety classification & learning model: defining & classifying potential serious injuries & fatalities. *Professional Safety, 69*(01), 18–26.

Hallowell, M. R., Hinze, J. W., Baud, K. C., & Wehle, A. (2013). Proactive construction safety control: measuring, monitoring, and responding to safety leading indicators. *Journal of construction engineering and management, 139*(10), Article 04013010. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000730

Hallowell, M. R., Alexander, D., & Gambatese, J. A. (2017). Energy-based safety risk assessment: does magnitude and intensity of energy predict injury severity? *Construction management and economics, 35*(1–2), 64–77.

Hallowell, M., Quashne, M., Salas, R., MacLean, B., & Quinn, E. (2021). The statistical invalidity of TRIR as a measure of safety performance. *Professional Safety, 66*(04), 28–34.

Hinze, J., Thurman, S., & Wehle, A. (2013). Leading indicators of construction safety performance. *Safety science, 51*(1), 23–28. https://doi.org/10.1016/j.ssci.2012.05.016

Hollnagel, E. (2014) Safety-I and Safety-II: The past and future of safety management. Ashgate Publishing Company.

Hoonakker, P., Loushine, T., Carayon, P., Kallman, J., Kapp, A., & Smith, M. J. (2005). The effect of safety initiatives on safety performance: a longitudinal study. *Applied ergonomics, 36*(4), 461–469.

Janicak, C. A. (2009). *Safety metrics: Tools and techniques for measuring safety performance*. Government Institutes.

Jazayeri, E., & Dadi, G. B. (2017). Construction safety management systems and methods of safety performance measurement: A review. *Journal of Safety Engineering, 6*(2), 15–28.

Kahane, C. J. (2015). Lives saved by vehicle safety technologies and associated Federal Motor Vehicle Safety Standards, 1960 to 2012 – Passenger cars and LTVs – With reviews of 26 FMVSS and the effectiveness of their associated safety technologies in reducing fatalities, injuries, and crashes. (Report No. DOT HS 812 069). Washington, DC: National Highway Traffic Safety Administration.

Kang, Y., Siddiqui, S., Suk, S. J., Chi, S., & Kim, C. (2017). Trends of fall accidents in the US construction industry. *Journal of Construction Engineering and Management, 143*(8), Article 04017043.

Kjellén, U. (2009). The safety measurement problem revisited. *Safety Science, 47*(4), 486–489.

Kjellén, U. (2023). Preventing fatal accidents in construction through the management of barriers. *Heliyon, 9*(11).

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine, 15*(2), 155–163.

Krause, T. R., and Murray, G. (2012, June). On the prevention of serious injuries and fatalities. In *ASSE Professional Development Conference and Exposition* (pp. ASSE-12). ASSE.

Laitinen, H., Marjamäki, M., & Päivärinta, K. (1999). The validity of the TR safety observation method on building construction. *Accident Analysis & Prevention, 31*(5), 463–472.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. biometrics, 159-174.

Lingard, H., Hallowell, M., Salas, R., & Pirzadeh, P. (2017). Leading or lagging? Temporal analysis of safety indicators on a large infrastructure construction project. *Safety science, 91*, 206–220. https://doi.org/10.1016/j.ssci.2016.08.020

Lofquist, E. A. (2010). The art of measuring nothing: The paradox of measuring safety in a changing civil aviation industry using traditional safety metrics. *Safety science, 48*(10), 1520–1529.

Manuele, F. A. (2005, June). The challenge of preventing serious injuries. In *ASSE Professional Development Conference and Exposition* (pp. ASSE-05). ASSE.

Martin, D. K., & Black, A. (2015). Preventing serious injuries and fatalities: Study reveals precursors and paradigms. *Professional Safety, 60*(09), 35–43.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia medica, 22*(3), 276–282.

McVittie, D., Banikin, H., & Brocklebank, W. (1997). The effects of firm size on injury frequency in construction. *Safety Science, 27*(1), 19–23.

Memarian, B., Brooks, S. B., & Le, J. C. (2023). Obstacles and solutions to implementing job hazard analysis in construction: a case study. *International Journal of Construction Education and Research, 19*(2), 187–198.

Oguz Erkal, E. D., & Hallowell, M. (2023). Moving beyond TRIR: measuring & monitoring safety performance with high-energy control assessments. *Professional Safety, 68*(05), 26–35.

Oguz Erkal, E. D., Hallowell, M. R., & Bhandari, S. (2023). Formal evaluation of construction safety performance metrics and a case for a balanced approach. *Journal of safety research, 85*, 380–390.

Oguz Erkal, E. D., Hallowell, M. R., Ghriss, A., & Bhandari, S. (2024). Predicting serious injury and fatality exposure using machine learning in construction projects. *Journal of Construction Engineering Management, 150*(3), Article 04023169.

OSHA. (Occupational Safety and Health Administration). (n.d.). Injury Tracking Application (ITA) Information. https://www.osha.gov/injuryreporting.

Oswald, D. (2020). Safety indicators: Questioning the quantitative dominance. *Construction Management and Economics, 38*(1), 11–17.

Oswald, D., Zhang, R. P., Lingard, H., Pirzadeh, P., & Le, T. (2018). The use and abuse of safety indicators in construction. *Engineering, construction and architectural management, 25*(9), 1188–1209.

Petersen, D. (1998). What measures should we use, and why? *Professional Safety, 43*(10), 37.

Provan, D. J., Woods, D. D., Dekker, S. W., & Rae, A. J. (2020). Safety II professionals: How resilience engineering can transform safety practice. *Reliability Engineering & System Safety, 195*, Article 106740.

Raheemy, Y., Sherratt, F., & Hallowell, M. R. (2025). What is safety? contemporary definitions and interpretations across North America. *Safety Science, 185*, Article 106798.

Rausand, M., and Haugen, S. (2020). *Risk Assessment – Theory, Methods, and Applications (2nd Edition)*. John Wiley and Sons.

Reiman, T., & Pietikäinen, E. (2012). Leading indicators of system safety–monitoring and driving the organizational safety potential. *Safety science, 50*(10), 1993–2000.

Roughton, J. E. (2002). The Benefits of Job Hazard Analysis. *OSHA 2002 Recordkeeping Simplified*, 170-184.

Roughton, J., and Crutchfield, N. (2015). Job hazard analysis: A guide for voluntary compliance and beyond. *Elsevier Science & Technology*.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical therapy, 85*(3), 257–268.

Sun, S. (2011). Meta-analysis of Cohen's kappa. *Health Services and Outcomes Research Methodology, 11*(3), 145–163.

Swartz, G. (2002). Job hazard analysis. *Professional Safety, 47*(11), 27.

Taubitz, M. A. (2018). PTD before risk assessment: A historical perspective. *Professional Safety, 63*(11), 26–35.

Vandeskog, B. (2024) Safety is the preservation of value. *Journal of Safety Research*, p. S0022437524000185. Available at: 10.1016/j.jsr.2024.02.004.

Versteeg, K., Bigelow, P., Dale, A. M., & Chaurasia, A. (2019). Utilizing construction safety leading and lagging indicators to measure project safety performance: a case study. *Safety Science, 120*, 411–421.

Zheng, W., Shuai, J., & Shan, K. (2017). The energy source based job safety analysis and application in the project. *Safety science, 93*, 9–15.

Zou, Y., Tarko, A. P., Chen, E., & Romero, M. A. (2014). Effectiveness of cable barriers, guardrails, and concrete barrier walls in reducing the risk of injury. *Accident Analysis & Prevention, 72*, 55–65.

Arnaldo Bayona Dr. Arnaldo Bayona is an Assistant Professor in the Department of Engineering Technologies, Safety, and Construction at Central Washington University. He holds a BS and a PhD in Civil Engineering, and an MS in Construction Engineering and Management. Before his academic career, he worked as a construction management intern and project engineer on residential and infrastructure projects. He also served as a contract administrator and district director for Ecuador's Ministry of Transportation and Public Works. His work focuses on advancing safety through applied research, with the goal of creating safer work environments free from serious injuries and fatalities.

Matthew R. Hallowell Dr. Matthew Hallowell is a President's Teaching Scholar and Endowed Professor of Construction Engineering at the University of Colorado Boulder. He earned a BS and MS in Civil Engineering and a PhD with a focus on Construction Engineering and Occupational Safety and Health. Before his academic career, he worked in construction as a laborer, project engineer, and quality inspector. Dr. Hallowell specializes in construction safety research and has pioneered leading indicators, hazard recognition programs, safety risk assessment, and precursor analysis. He has published over 100 peer-reviewed papers and has overseen over $3 million in external funding. For his research, he has received the National Science Foundation CAREER Award and the Construction Industry Institute Outstanding Researcher Award.

Siddharth Bhandari Dr. Siddharth Bhandari is the Associate Director of Research of the Construction Safety Research Alliance. He received his Ph.D., in Civil Engineering in 2019, M.S. in Civil Engineering in 2017, and B.S. in Civil Engineering in 2014 from the University of Colorado at Boulder. Siddharth's thesis focused on improving the current construction safety training techniques by leveraging behavioral psychology and adult learning principles critical to improving workers' hazard identification, risk assessment, and decision-making skills. For his research, Siddharth has received the best paper award at the 2015 5th International/11th Construction Specialty Conference in Vancouver, Canada. Siddharth's research expertise exists at the nexus of human factors engineering, behavioral psychology, safety climate, and adult learning to improve safety and productivity on construction sites. He has also received the ELECTRI International's Early Career Award 2019 to develop an empirically validated tool to measure situational awareness, human factors performance, and safety climate in real-time on construction sites

Mohammad Yaqoob Raheemy With over five years of diverse experience as a skilled project manager and planning engineer with construction companies and consultancies in the UK and Afghanistan, Yaqoob Raheemy brings a wealth of expertise to the academic realm. Yaqoob's academic journey includes a Master's degree in Construction Project Management from Birmingham City University in England, UK, and an undergraduate degree in Civil Engineering from Kabul University in Kabul, Afghanistan. Yaqoob also holds esteemed certifications such as Project Management Professional (PMP) from the Project Management Institute, as well as being a full member of the Association for Project Management (MAPM) in the UK. Currently, Yaqoob is engaged as a graduate research assistant, contributing to the "High Energy: Controlling the Uncontrollable" research project. This project will explore why direct controls are sometimes challenging; how to determine if an alternative work plan is sufficient; and how to continuously monitor the state of high-energy when workers are exposed.