Journal of Safety Research 98 (2026) 642–656

Contents lists available at ScienceDirect

Journal of Safety Research

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

Injury rates tell an incomplete story, but a more complete measurement narrative may be possible

Kevin Geddert *, Sidney Dekker, Drew Rae

Griffith University – Nathan Campus, 170 Kessls Road, Nathan QLD 4111, Australia

ARTICLE INFO

Keywords:
Safety
Measurement
Injuries
Narrative
Severity

ABSTRACT

Introduction: Despite years of research challenging their validity, injury rate metrics continue to significantly influence decision-making in many organizations. In this paper, we argue that both the source of this problem and a solution rest in understanding the narrative power of these metrics. We explore this idea by developing and applying a set of alternative injury measurement tools specifically designed to provide more comprehensive and compelling narratives of injury performance. Comparing alternative measures with traditional stand-alone injury rates demonstrates how the overall narrative of injury performance changes through metric design. Our overall goal is not to fix the statistical validity of the metrics, but to improve safety decision-making through more complex discourse. *Method:* The analysis in this paper rests on three large organization case studies. In each case, study data from the organization was presented back to senior executives using: the injury rate alone; the average injury severity; and a weighted combination of rate and severity known as SAIF. The resultant conversations were then analyzed, supplemented by individual interviews with conversation participants. *Results:* This research highlights the importance of presenting injury rate and severity measures together and demonstrating their interplay, rather than as stand-alone metrics. Four guidelines are proposed for improving narratives and conversations based on injury metrics. The four guidelines are: pairing measures, continuous variability, a simple communication mechanism, and highlighting differences. *Conclusions and Practical Applications:* Together, these guidelines enable users of injury measures to generate more dynamic discussions from their metrics.

1. The narrative of safety performance as presented by our measures

1.1. The injury rate as a narrative tool, not a statistical tool

The injury rate – also referred to as Total Recordable Injury Frequency (TRIF) or Total Recordable injury rate (TRIR) – is calculated as the count of reported injuries classified as "recordable" per million or per 200,000 worked hours, depending on jurisdiction. It was originally conceived as a statistical tool (Bureau of Labor Statistics, 1916; Geddert, 2025b), and discussions of its merit and drawbacks are commonly framed in terms of its merit as a statistic.

Criticism of injury rates has been ongoing for decades across multiple fronts. Injury rates contribute to the pressure to under-report (Probst & Graso, 2013). Injury rates take attention away from major process risks (Hopkins, 2009). Injury focus can distract from the true work of safety (Dekker, 2017b). The injury rate is statistically meaningless in certain use cases (Hallowell et al., 2021). Modern safety programs ought to focus on "what went right" (Conklin, 2019; Dekker, 2017a; Hollnagel E, 2015; Provan et al., 2020). Given these challenges, one might expect a decline in the use of injury rates as a key measure of safety performance.

However, injury rate use continues. When defining "Safety" the most prominent response was the prevention of injury, incidents, and accidents (Raheemy et al., 2025). A literature review found a majority of studies defined safety performance as "the number of accidents and incidents or lagging safety indicators such as Lost-Time Injury Rate (LTIR) and Total Recordable Injury Rate (TRIR)" (Orvik et al., 2026). Major organizations still report injury rates in board reports as a representation of their safety performance (Coca-Cola, 2021; Shell, 2022). The continued use of injury rates despite research criticism requires an explanation.

A recent paper in this journal (Pomeroy & Pilbeam, 2025) suggested that metrics such as the injury rate hold meaning well beyond their use in spreadsheets. Pomeroy and Pilbeam used semiotic analysis to highlight 17 ways injury rates are interpreted within organizations, with only one of these meanings directly reliant on measurement validity.

* Corresponding author at: Safety Science Innovation Lab, Griffith University – Nathan Campus, 170 Kessls Road, Nathan QLD, 4111, Australia.
E-mail address: Kevin.geddert@griffithuni.edu.au (K. Geddert).

https://doi.org/10.1016/j.jsr.2026.07.018
Received 22 March 2026; Received in revised form 18 May 2026; Accepted 31 July 2026
Available online 1 September 2026

Their findings demonstrate the utility of examining the role of injury rates as signs, rather than as statistics. If injury rates are not interpreted as statistical measures, that may explain why criticism of them based on statistical validity is not persuasive.

This paper builds upon the work of Pomeroy and Pilbeam by examining the way injury rates create meanings through narrative generation. The stand-alone injury rate allows users to tell a story of improved performance when rates decline or worsening performance when rates increase. As a lagging measure, injuries (or their absence) are often viewed as the ultimate outcome of safety activities (Fu et al., 2020). The narrative power of injury rates lies in their ability to lead users to believe that their leading activities yielded the results they hoped for, even when causation cannot be established.

This paper examines the role of the injury rate as a narrative-generating tool by presenting three case studies. Injury rates from each organization were analyzed alongside alternative metrics and then presented back to the organization for reaction and comment. The alternative metrics were purposefully designed to create dynamic narratives rather than to statistically evaluate injury rates. This process asks how multiple measures change the narrative around injuries, whether those changes are meaningful, and which measurement features enable more dynamic results.

Three large organizations participated in this research, creating case studies that cover multiple industries, countries, and risk profiles. Cumulatively, the research participants contributed 947 million exposure-hours of data and 5,919 recordable injury events.

2. Three case studies using narrative-designed injury assessment tools

2.1. Study design

The study design required three steps:

1. A design for a set of narrative-based injury measures
2. Preparation of participant data with different approaches to injury measurement
3. Presentation to the participants and collation of feedback

Participation in this research was done anonymously to avoid market-perception risk. All three case studies are the actual data results for the organizations, with identifying names and references removed. The industries covered by the three organizations are diverse, including Telecommunications, Transportation, Manufacturing, onshore and offshore Energy Production, and Personnel Service Businesses.

Each participant provided data on exposure hours and specific injury metadata, without any identifiable personal information. The details of each participant's contribution are documented in their case study results.

2.2. Alternative measure design

2.2.1. Severity as a basis for assessing the injury rate

Injury rates require three things to generate: events, hours, and a definition of the severity of the included events. Severity and rate of events are both significant contributors to injury performance and have been since the measure's inception (Bureau of Labor Statistics, 1916; Geddert, 2025b). Injury rates are fundamentally a normalized measure of events of a specific severity. Severity is commonly assessed using alternative injury rates, such as Days Away Restricted or Transferred (DART rates), Lost Time (LTI) rate, or Severe Injury and Fatality (SIF) rates. These are the same measures as a recordable injury rate, but with smaller datasets. Assessing severity this way generates the same better-or-worse narrative because injury severity and rate are functionally measured with the same tool.

Moving between different versions of the same measurement tool (TRIF, LTI, DART, SIF, etc.) may change the focus area, but the narrative doesn't become more dynamic. Safety can be suggested to improve or diminish due to a shift in focus for a period, but this is still a better-or-worse narrative. Stand-alone rates make it difficult to tell a simple or compelling narrative about how rates and severity both change over time, or how they differ across groupings of organizations or failure mechanisms.

2.2.2. Designing a simple severity narrative

Creating a simple narrative tool around severity is challenging, as injury severity is fundamentally complex. Multiple concepts govern our understanding of injury severity, including medical treatment, restricted work time, lost work time, and lasting impacts on quality of life. One severity measure addresses this with the average energy of failure mechanisms, which provides a level of accuracy for the mathematics (Hallowell & Erkal, 2024). The narrative challenge is that the average person doesn't think about injury severity in terms of energy output, but rather impact on human life (Smith & Jones, 2014). Any single variable (such as energy output or the lost time used in the original severity rate) will favor one category of events while not necessarily being representative of others. Severity becomes so complex to measure that a simple narrative that connects with everyday people seems impossible.

Injury measurement can use the Wind Chill model to create a simple narrative. Paul A. Siple coined the term " wind chill " in his unpublished thesis, "Adaptation of the explorer of the climate of Antarctic"(Siple & Passel, 1945). The results of his research were presented in units of kilogram calories per hour per square meter. The presentation of the mathematically complex concept of Wind Chill, as a relative temperature, would be credited to the US Army Alaska Command, which printed copies of the wind chill relative temperature table to give to soldiers (Consolazio, 1963; Eagan, 1964). Eagan would describe the usefulness of the relative temperature.

> *Persons unfamiliar with the wind chill index and with little formal education can, nevertheless, understand the concept of an equivalent temperature under calm conditions.* p.151 (Eagan, 1964).

Using relative temperature, soldiers could dress appropriately for the weather without needing to understand the complexities of a mathematically sound wind chill index. Further research would correct for proper wind speed assumptions and freezing effects specific to the face (Osczevski & Bluestein, 2005; Osczevski, 1995), but results would be presented as relative temperatures to make the mathematics usable to everyday people.

Two concepts from wind chill are of particular significance for safety:

- Defining a complex measure, from a simple one, can make the complex measure usable without a deep understanding of the complexity
- The usability of the measure is more important than absolute accuracy.

Translating a relative injury rate is not enough; a measure needs to represent something meaningful to the average person. An injury-specific solution to this challenge is the concept of a benchmark injury, which serves as the standard by which all others are evaluated. The benchmark injury defined in this research method is an event that results in two weeks of lost time. This was selected for several reasons; the most substantial is that a two-week LTI balances a definition of a significant amount of lost time with the ability to generate the measure in a reasonable timeframe to be considered trustworthy. All other injuries can now be mathematically assessed based on whether a reasonable person might consider an injury better or worse than two weeks lost time.

An injury resulting in a two-week LTI is not more or less important than any other injury; the benchmark is a communication tool. Those unfamiliar with the underlying complexity of severity can nonetheless understand that the measure is relative to two weeks lost time. The injury rate is the number of people injured per unit of time, and the severity rate is the same thing adjusted for whether the person was off for two weeks. Calculating such a measure would require complex calculations, but understanding that complexity is not necessary for the average user. The severity rate can be explained to someone with a limited understanding of or interest in the complexity of severity, in terms they find intuitively usable.

Injury rate: The number of people hurt / exposure hours.

Severity Rate: The relative number of people with a two-week LTI / exposure.

hours.

2.2.3. The math behind the SAIF rate

The Severity Adjusted Injury Frequency (SAIF Rate) is defined as follows:

1. The SAIF Rate modifies the numerator of the Injury Frequency
2. An injury with two weeks lost time is the benchmark because it is mathematically equal for both TRIF and SAIF (a 1 in both cases). It is the crossover point between the measures.
3. All other injuries are given a base score depending on their classification, with a modifier in each classification.
4. Base scores for each category use a base 3 logarithm, so each category is 3 times more or less significant than the surrounding category.
5. The modifiers in each category adjust the resulting SAIF numerator based on severity definitions in that category

Table 1 describes the base scores for each category and the modifiers within each category.

Each event is evaluated based on the category score and category modifier. The category modifiers are designed specifically to recognize differences within each category. They also avoid categorical leaps in value, which might motivate users toward under-reporting by either exclusion or recategorization. The category modifiers for Lost Time and Restricted Work are: if an event is under 14 days, it is scored as a medical treatment case + percent of the remaining category score. A SIF is defined in accordance with the Life Model (Bayona et al., 2023), with the modifier being a fatality, which is doubled. A limited medical treatment is worth 0.065 and is defined as someone whose only medical treatment, beyond first aid, was a scan or prescription.

The Severity Adjusted Injury Frequency is calculated using the same methodology as injury frequency, with a base of either 1,000,000 or 200,000 exposure-hours, depending on the region.

TRIF = (Count of recorded incidents as 1) / (1,000,000 or 200,000 h depending on region) *.

SAIF= (Sum of all incident SAIF scores) / (1,000,000 or 200,000 h depending on region) *.

*These rates are typically calculated on a 12-month rolling basis to smooth monthly variations. The denominator must be the same for both measures; the options are presented because of differences across certain world regions. The SAIF rate was trialed first with the coal seam gas industry in Queensland, Australia, and published in a conference case study paper (Geddert, 2025a).

Table 1
SAIF Category Calculations.

Category	Base	Modifier	Math
Serious Injury or Fatality	3	Fatality	Doubled
Lost Time	1	Less than 14 days	0.11 + 0.89(Days/14)
Restricted Work	0.33	Less than 14 days	0.11 + 0.22(Days/14)
Medical Treatment	0.11	Limited Medical Treatment	Halved

2.3. Data preparation and analysis design

2.3.1. Preparing results

Similar charts were prepared for the three anonymous participants. The results use the following measures:

- **Total Recordable Injuries**: The sum of recordable injuries, with all equal to 1
- **Total Recordable Injury Frequency**: Total Recordable Injuries per million manhours on a 12-month rolling basis.
- **Severity Adjusted Injuries**: The sum of injuries with all events adjusted according to the SAIF methodology
- **Severity Adjusted Injury Frequency**: Cumulative SAIF scores per million manhours on a 12-month rolling basis.
- **Average SAIF score**: The sum of Severity Adjusted Injuries divided by the sum of Total Recordable Injuries

2.3.2. Analyzing results

The first step in this research identified trends of six or more months in the stand-alone injury rate. Individual period variations were not reviewed in detail. This approach to evaluating injury rates is similar to how companies review their own injury rates. The data were then re-evaluated with the alternate measures (a matched set of injury rate, SAIF, and Average SAIF). Areas that were highlighted as trends with a stand-alone injury rate were subsequently assessed to determine whether their original narrative held or a new narrative emerged using the alternative measurement approach.

3. Case study results

3.1. Participant 1 – A case study

The first participant in this study operates across multiple industries; the organization contributed data on 4,626 recordable injuries and 757 million exposure hours over 13 years. The data included individual events and hours, broken down by working groups, and injuries classified by mechanism and body part.

3.1.1. Participant 1 – Overall story

The narrative from TRIF alone was that the organization had seen a downward trend over the past 13 years. More recent trends include a rise in the injury rate from Jan 2019 (4.48) through to Jan 2020 (7.24). This rise in injury rate was followed by a period of slight decline between Jan 2020 (7.24) and May 2021 (6.3). A turning point in May 2021 culminated in the company's lowest TRIF score in September of 2023 (TRIF 3.23). Following that period the injury rate essentially stagnated, varying slightly upward through 2024 and into 2025. These trends served as the baseline for evaluating the narratives generated by an alternative measurement approach. (See Fig. 1)

The narrative of steadily improving safety changed when viewed with a Severity Adjusted Injury Frequency (SAIF). The earlier period in 2012 was not necessarily the "worst" safety performance. The period with the highest SAIF was March of 2020. From Jan 2012 through March 2020 the organization declined from 16.66 to 7.24 TRIF (−56.5%); the average event SAIF score increased from 0.23 to 0.55 over the same period (+139%). That trend would continue into 2024, where a lowering of TRIF from 7.24 down to its low of 3.23 was mirrored by an additional increase in average event severity from 0.55 to 0.74. In the same period where the long-term trend was an 80% reduction in injury rates (16.66 down to 3.23), the average severity of events would more than triple (from 0.23 to 0.74). (See Fig. 2)

Between Jan 2019 and May 2021, there was a period of relative stability in the average severity of events. Average event scores stayed between 0.50 and 0.55 through both the climb and the slight decline

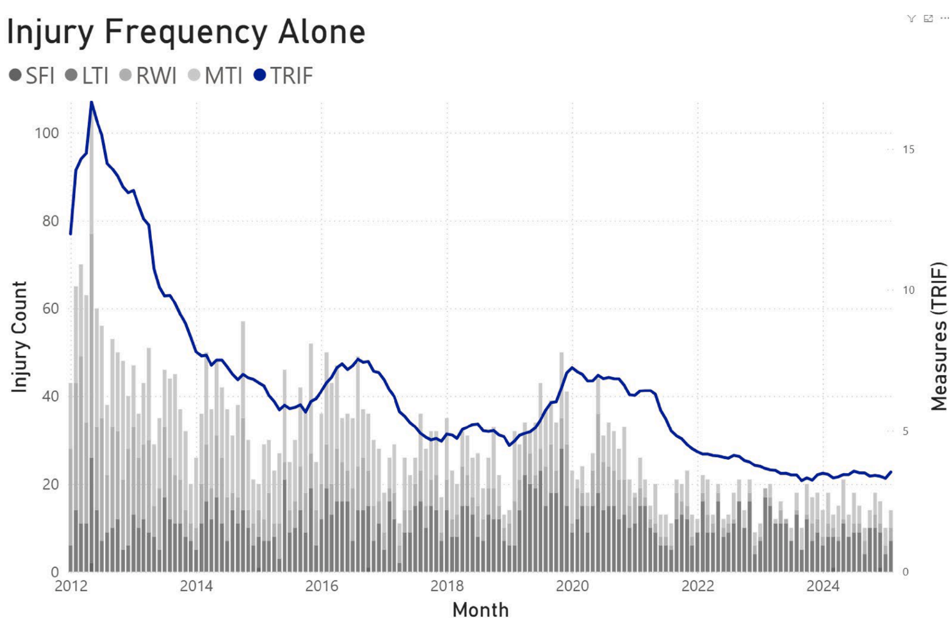

Fig. 1. Participant 1 – Full Company TRIF Alone.

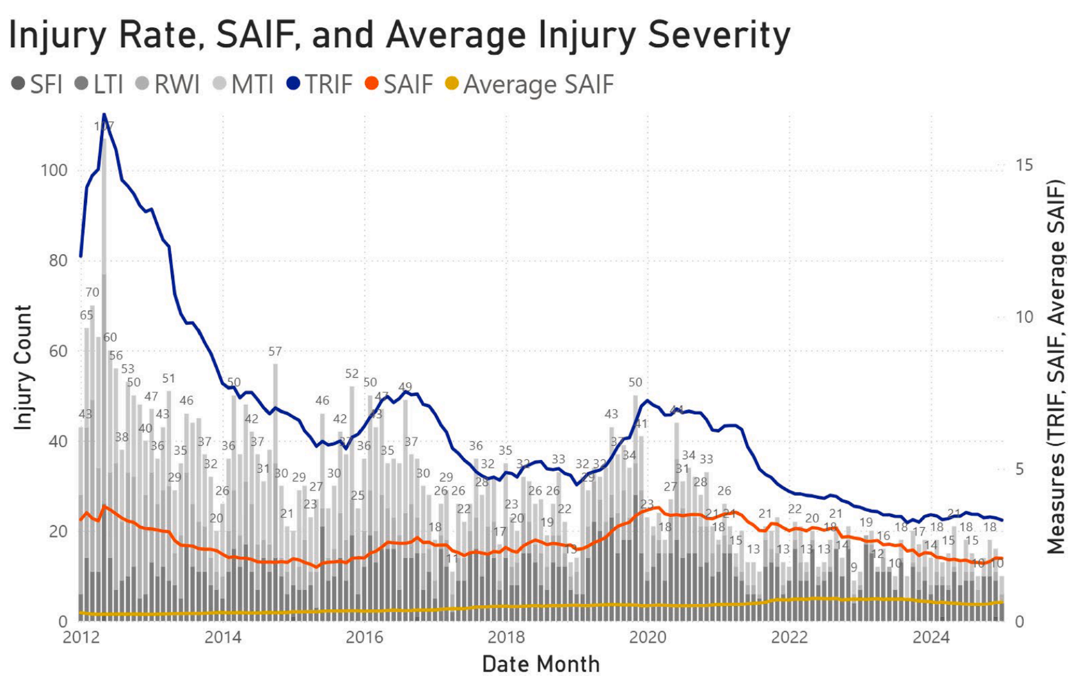

Fig. 2. Participant 1 – Full Company TRIF, SAIF and Average SAIF Rates.

periods in TRIF. This suggests that the change in injury rate was due to an increase in the volume of events with a similar profile. For these periods, the changes in TRIF reflected changes in rates of similar events.

In terms of a changing narrative, the long-term reduction in TRIF was directly offset by a long-term increase in the average severity of events. Those two pieces of information together tell us that reduced injury rates were more reflective of changes in the rate of low-severity events than of higher-severity events. The data cannot determine whether this was because lower-severity events were better controlled or simply were not classified as recordable injuries in a similar manner across periods.

An additional narrative was significant in the alternative analysis. The injury rate stagnated in Sept 2023 and even climbed slightly. With stand-alone injury rates this suggested that a lower boundary may have been reached. The alternative analysis demonstrated that the Severity-Adjusted Injury Rate continued to decline. The average severity of events started to come down for the first time between April 2023 (0.74) and a low of 0.55 in September 2024. This period would see the first material drop in the average severity of events in 12 years, suggesting that a material change in injury performance was taking place while injury rates held flat.

The lowest overall SAIF score was achieved in May of 2015 (1.76), and a close second was achieved in October 2024 (1.88). The period previously thought to be stagnating or even deteriorating in injury rate performance represented one of the best cumulative improvements in both event severity and volume in the organization's history under an alternative approach to injury measurement.

3.1.2. Participant 1 – Comparing Internal businesses

Participant 1 has four key businesses that operate across different industries and have distinct risk profiles. Common practice was to compare the four businesses using TRIF measures. The “Other” business was not evaluated, as it consists primarily of support functions and has a lower risk profile than the operating businesses. (See Table 2)

Participants first compared their businesses using the stand-alone TRIF metric, then were given the full Table 1. The common narrative within the business was to rank the businesses from best to worst safety performance based on their TRIF scores (making business 2 the worst, and business 3 the best). (See Fig. 3)

Adding the alternative measures significantly changed the order of perceived business performance. Business 3 still had the lowest SAIF rate (it has a lower risk profile). The surprising finding was that Business 2 would have the next lowest SAIF rate, because it had the lowest average event score. (See Fig. 4)

Businesses 1 and 2 demonstrated a considerable difference in the overall severity of events. Business 2 had a consistently higher TRIF, while Business 1 had a consistently higher average event severity. In 2022, Business 1 rose to an average event severity of 0.87, bringing its SAIF line close to its TRIF line. In both businesses, the period from 2020 through 2023 saw a decrease in TRIF while the Average event increased. In both businesses, the 2024 period saw a levelling off of TRIF, but an ongoing reduction in the average severity of events, resulting in a lower SAIF as TRIF held flat or increased.

An unexpected narrative event occurred during this research. In early 2024 a significant injury would affect Business 4. This event would have little to no impact on the TRIF score for the period, as TRIF treats all events as equal. The event would, however, drive significant movement in the SAIF rate, as it was serious. An unmoving TRIF was described as dishonoring the sense of safety experienced following a significant event. The movement in the SAIF rate was welcomed, as it reflected not only severity but also the sentiment following a significant incident.

3.1.3. Participant 1 – Mechanisms of injury

When comparing failure mechanisms, “Hitting Objects” was the largest contributor to the TRIF rate but ranked fourth in its contribution to SAIF. The most discussed narratives were those of lower-frequency events with higher-severity outcomes. Falls from Height, Bullying and Assault, and Motor Vehicle accidents had the highest average event severities. This finding was not considered surprising, but rather a visible reflection of the reality of failure mechanisms in the business. These results gave credibility to the measures, and safety work in the business, as they reflect the understood differences between failure mechanisms. Practitioners could use this analysis to demonstrate the higher average severity of certain low-frequency events, particularly those considered fatal risks warranting attention. (See Figs. 5 and 6)

The comparison of event volumes and severity was not uniformly distributed across businesses. This example, from Business 2, demonstrates that motor-vehicle injuries in that group have the most significant outcomes by a large margin, despite their low frequency. Other businesses saw greater averages in bullying and assault, or falls from height.

While not central to this research, one question should be addressed here: This assessment does not necessarily reflect the risk profile of a business. If a business were to have a significant risk that it had controlled completely, nothing would show up here. This assessment measures the rate and magnitude of risks that remain uncontrolled.

Table 2
Participant 1 TRIF, SAIF and Average SAIF – Jan 2025.

Business	TRIF	SAIF	Average SAIF	Total Hours
Business 1	3.50	2.39	0.68	301,047,920
Business 2	5.20	2.12	0.41	172,550,471
Business 3	1.49	0.78	0.52	203,488,031
Business 4	2.44	2.42	0.99	58,189,539
Other	1.22	1.22	1.00	11,790,427
Total	**3.31**	**2.05**	**0.62**	**747,066,388**

3.2. Participant 2 – A case study

Participant 2 is a large Australian company. They have a single core business with multiple locations, but most of their operation and data comes from two major locations. The data provided covers 6 years, 144 recordable injuries, and 46 million exposure hours.

3.2.1. Participant 2 – Overall story

The initial narratives from a stand-alone TRIF were that the organization experienced an increase in injuries from 2019 through the beginning of 2022, rising from a low of 0.92 to a high of 5.14. From the high point in June 2022, the organization's TRIF rate declined, reaching a low of 1.89 in May 2024. The latter part of 2024 saw the rate return upward to 2.59. (See Fig. 7)

The period between Jun 2020 and Jun 2021 saw a 37.6% increase in TRIF from 2.5 to 3.44; however, it also saw a corresponding 39.5% decrease in average severity from 0.43 to 0.26. During this period, the story told by TRIF was one of worsening performance; when coupled with the severity of events, there was little change in the relative number of people with two week lost time (the SAIF rate). (See Fig. 8)

The period from August 2021 through June 2022 saw TRIF rise from 3.7 to 5.14, while the average event severity remained between 0.32 and 0.34. This suggests that, during this period there were more events of a similar severity, which was reflected in SAIF Rate movement. The introduction of paired analysis enabled recognition of differences between a period in which TRIF increased while average severity declined, and a period in which TRIF grew materially with a similar event profile.

The period between March 2023 and May 2024 saw a drop in TRIF from 3.52 to 1.89 (a 46.3% reduction). That period would also see the average severity of events rise from 0.25 to 0.39 (a 56% increase). This period is like that observed in Participant 1. The drive to reduce the injury rate was successful, at the expense of increasing event severity. As with participant 1, the increasing average severity of events indicates that the material decrease in TRIF was more effective at reducing lower-severity outcomes. The data cannot tell us whether underreporting was part of the equation, only that lower-severity events were reduced at a greater rate than more severe events.

During the period from May 2024 to December 2024, TRIF increased from 1.89 to 2.59 (a 37% increase), and the average event severity decreased from 0.39 to 0.28 (a 28% decrease). The addition of more lower-severity events would see the SAIF rate continue to decline. As a weighted average, the SAIF Rate doesn’t experience the same swings that are evident in the injury rate when the proportion of lower-severity events either grows or shrinks.

3.2.2. Participant 2 – comparing Internal businesses

Participant 2 had data presented by locations. Two locations accounted for 93.8% of the provided injuries. (See Table 3)

These two businesses had similar TRIF performance at the end of the reporting period, with Business 2 6.5% above Business 1. Although the two businesses are similar in terms of TRIF, Business 2′s SAIF rate is considerably lower than Business 1. Business 2 had experienced fluctuations in TRIF while maintaining a consistently low average event severity. (See Fig. 9)

Since 2021, Business 2 had experienced a steady decline in TRIF rate. There was a slight increase in the average severity of events between the TRIF high in May of 2021 and June of 2022. During this period, Business 2 experienced a 43% decline in TRIF but a 31.4% increase in the average event severity (from 0.296 to 0.389). This resulted in a SAIF rate of 2.68 over the period. From Jun 2022 onward, the average injury severity dropped to 0.198. This changed the narrative to suggest that the real

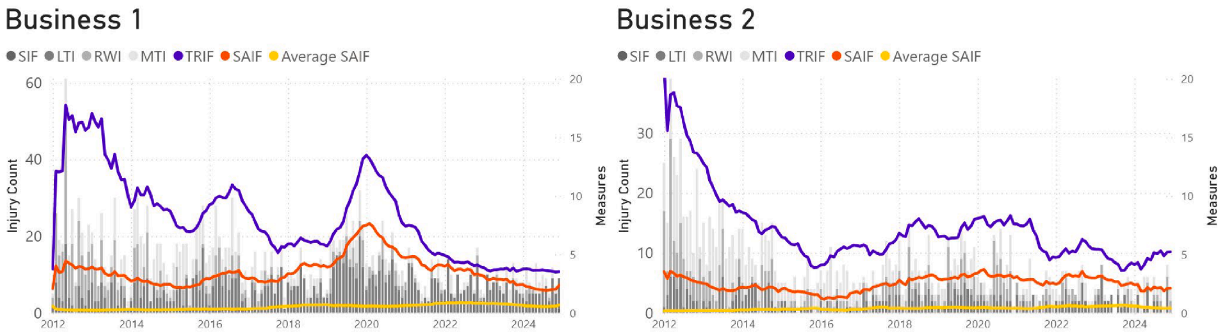

Fig. 3. Participant 1 – Business 1 and Business 2 Comparison.

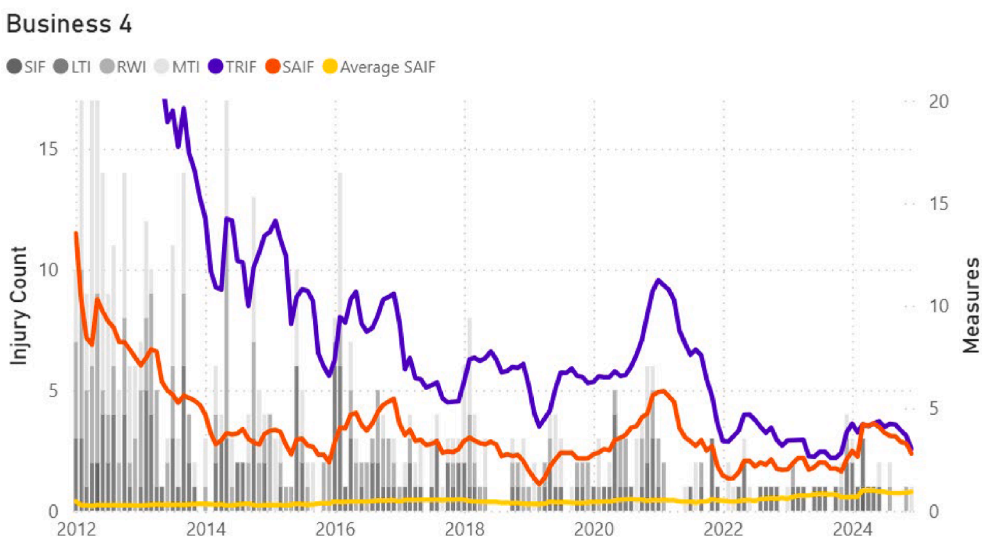

Fig. 4. Participant 1 – Business 4.

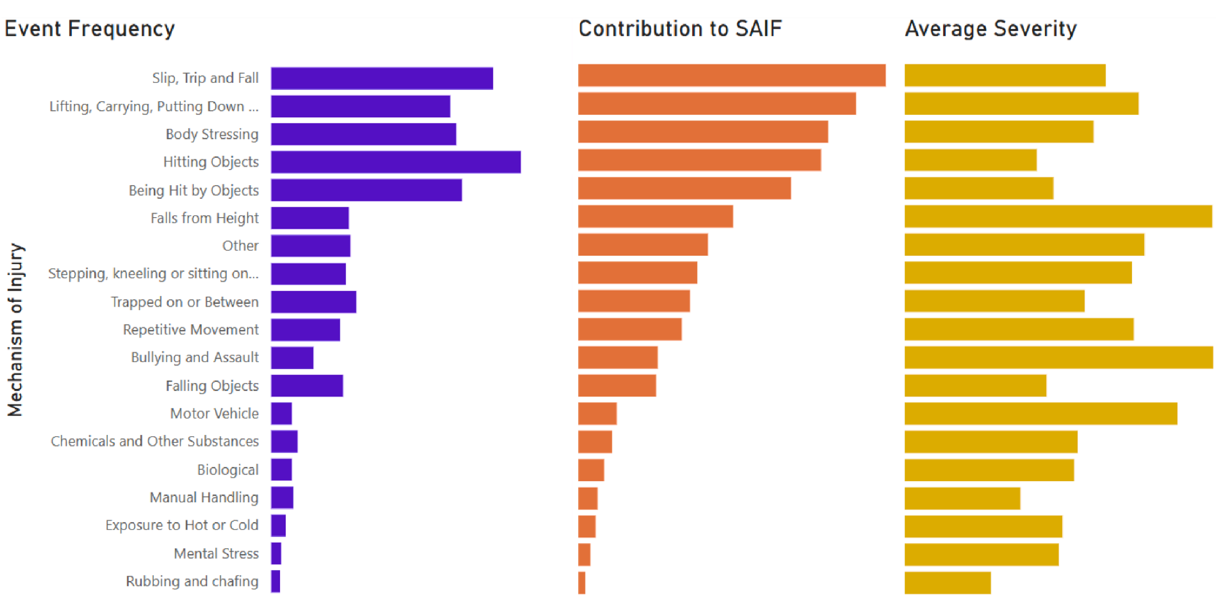

Fig. 5. Participant 1 – Frequency and Severity by Mechanism.

movement in Business 2 was not at the turn of TRIF in May of 2021, but the lowering of both TRIF and Average event severity in July of 2022. This combined reduction in both the rate and severity of events is reflected in the reduction of the SAIF as the weighted frequency rate.

For Business 1, the period from 2019 through to Feb 2021 would see TRIF fall to nearly zero, followed shortly by several significant incidents in 2021 that would see an increase in both TRIF and SAIF. Those events would roll off to lower TRIF and SAIF rates in November 2022. This

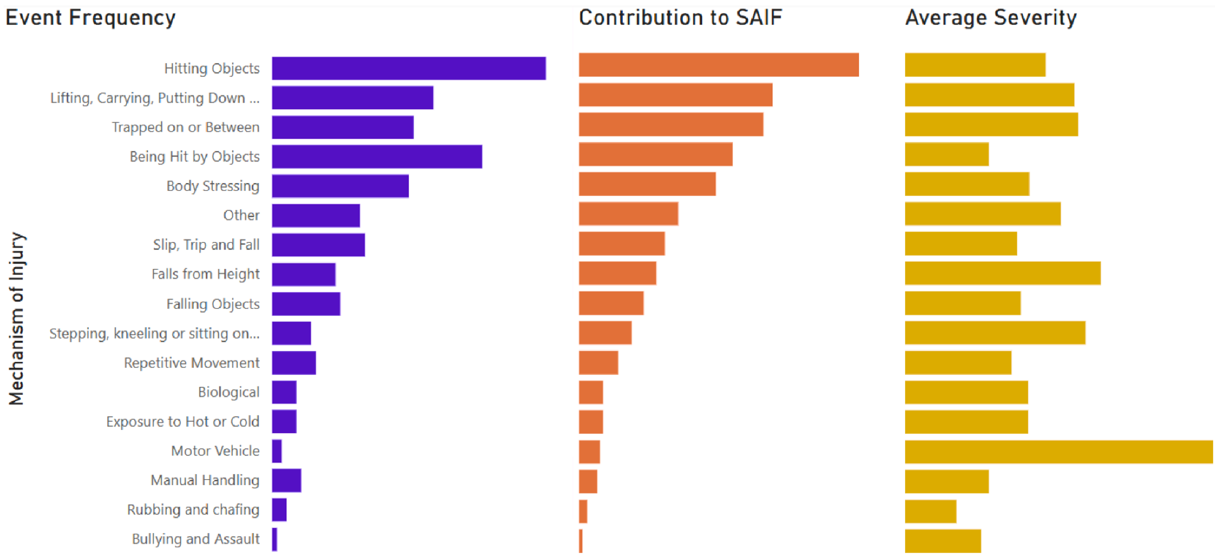

Fig. 6. Participant 1: Business 2 – Frequency and Severity by Mechanism.

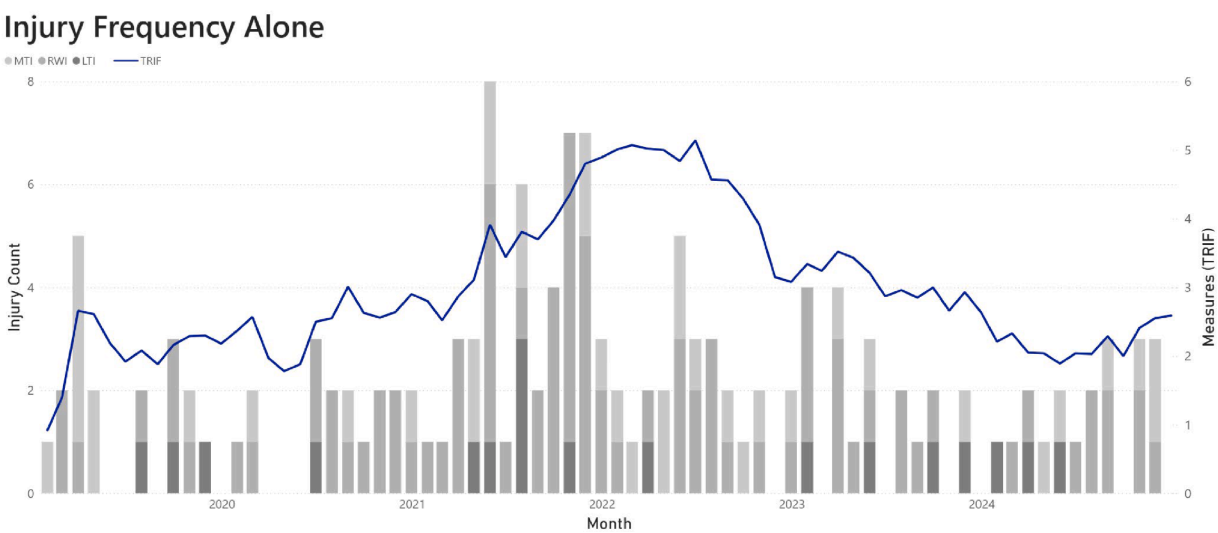

Fig. 7. Participant 2 – Full Company TRIF Alone.

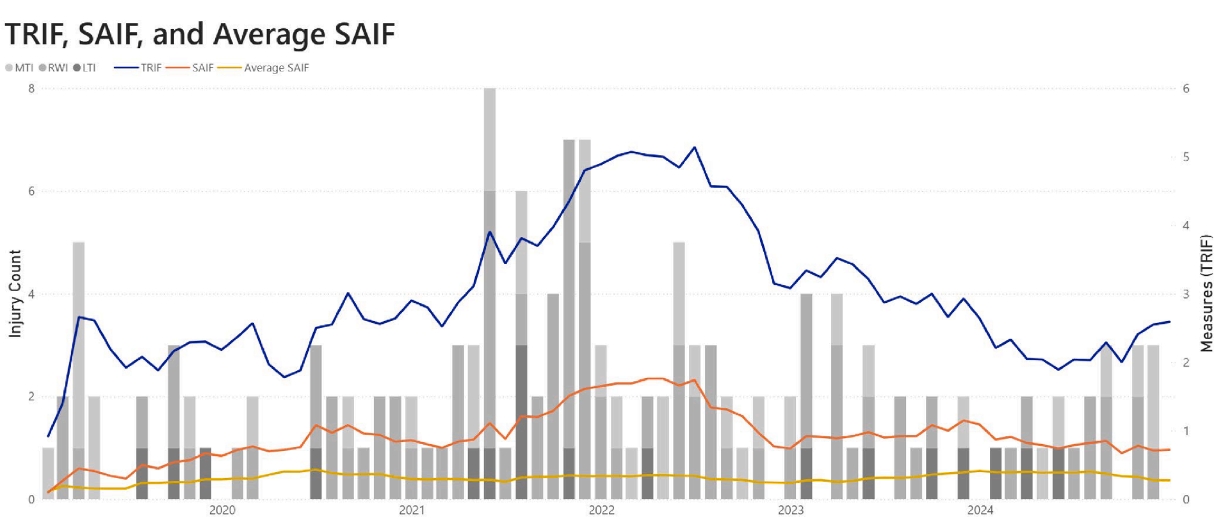

Fig. 8. Participant 2 – Full Company TRIF, SAIF and Average SAIF Rates.

Table 3
Participant 2 TRIF, SAIF and Average SAIF – Jan 202.

Business	TRIF	SAIF	Average SAIF	Total Hours
Business 1	3.07	1.18	0.38	15,026,040
Business 2	3.27	0.67	0.20	15,113,519
Total	**3.16**	**0.95**	**0.30**	**30,139,559**

turnabout would support the narrative that the elimination of TRIF may not indicate a controlled safety environment.

In Nov 2022 the injury rate was 4.42. The injury rate would eventually drop from 4.42 to 2.27; the average severity of events would increase from 0.195 to 0.511, resulting in a SAIF rate above the Nov 22 start (0.86 to 1.16). This once again demonstrates an inverse relationship between the rate of events and the average severity of those events, suggesting that a reduction in injury rate alone cannot be celebrated as improved injury performance without an understanding of the underlying severity of those events.

3.2.3. *Participant 2 – Mechanisms of injury*

Participant 2 demonstrated a pattern similar to Participant 1, with certain known failure mechanisms showing higher average severity. Participant 2 highlighted repetitive movement injuries as a high-severity outcome and muscular stress as a moderate-severity, high-frequency event. This analysis again helped confirm the rate and magnitude of various uncontrolled risks for a specific business activity and also highlighted events with different profiles of frequency versus potential outcomes. (See Fig. 10)

3.3. *Participant 3 – A case study*

The third participant is a multinational organization that provided two data sets. The first set of data was for their larger international reporting entity, which was collated quarterly (referred to as 3A). The quarterly dataset comprised 1,149 recordable injuries and 144 million exposure hours over 42 months. The second dataset was specific to their Australian operation and detailed individual events monthly (referred to as 3B). The Australia-specific dataset included 520 recordable injuries and 44.9 million exposure-hours. The data in 3B were a subset of those in 3A.

3.3.1. *Participant 3A – The overall organization story*

Ignoring the first three quarters (which are necessary to establish a

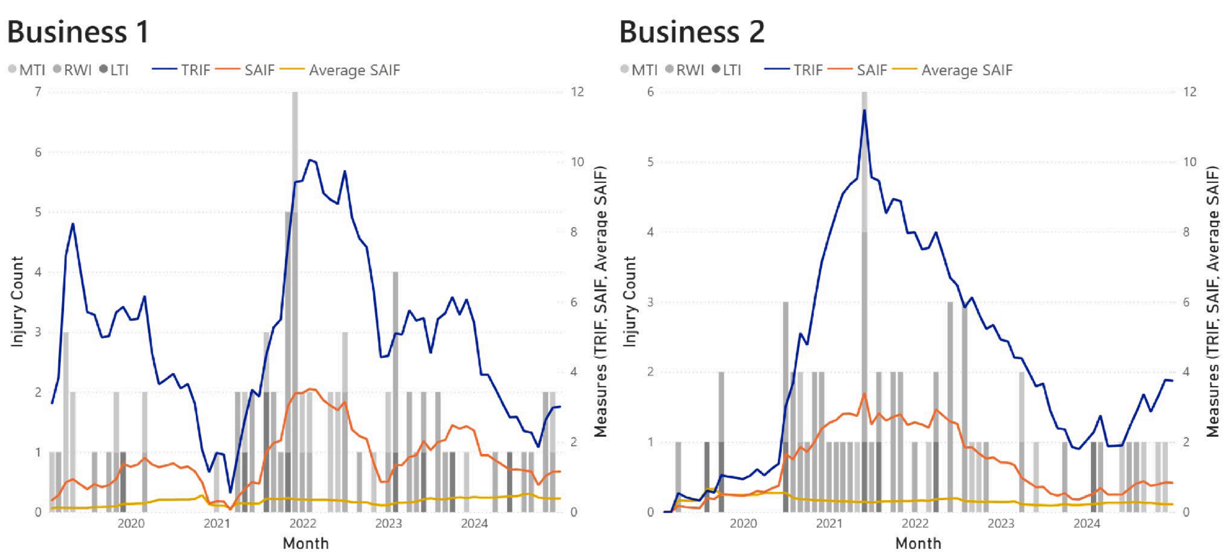

Fig. 9. Participant 2 – Business 1 & 2 Comparison.

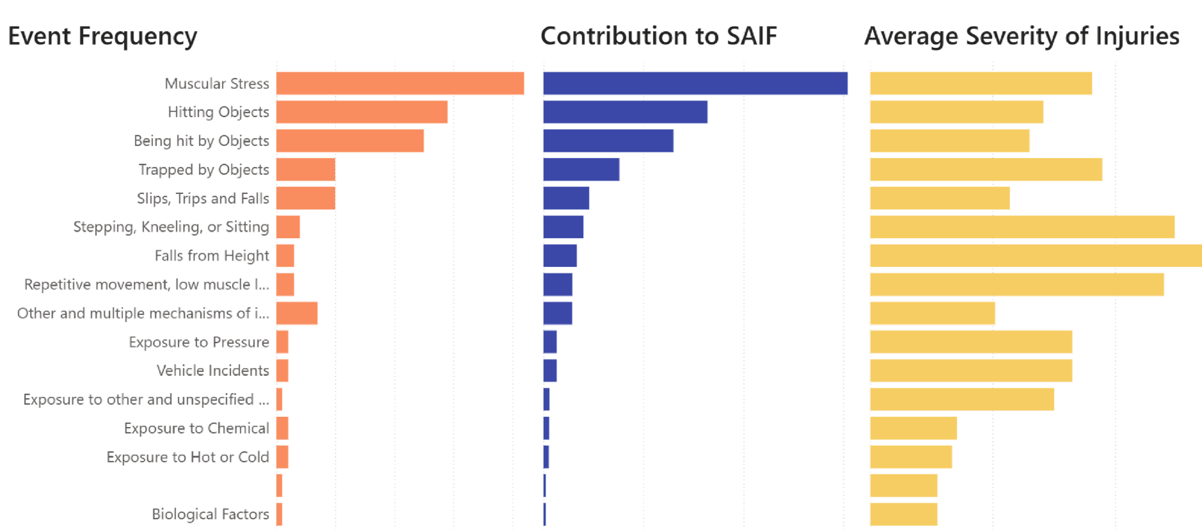

Fig. 10. Participant 2 – Frequency and Severity by Mechanism.

12-month rolling metric), the TRIF chart indicates that the organization experienced an increasing injury rate over two years, from 7.07 to 9.3. The growth was consistent, with only one quarter showing a decline before returning to the upward trend. (See Figs. 11 and 12)

The alternative analysis demonstrates that, at the full company level, the organization has experienced a 31.5% increase in injury rate over two years, with only a marginal change in average event severity (0.45 to 0.43; −4.4%). This organization experienced growth in TRIF and SAIF Rate with a similar event profile. The results from the international business units would challenge the validity of that narrative.

3.3.2. Participant 3A – Comparing Internal businesses

Participant 3A provided an international dataset. The data were prepared to compare TRIF, SAIF, and the average SAIF by country. The countries in this research are anonymized. Countries 1, 2, and 4 are individual operating countries. Country 3 is one main country with a few smaller ones amalgamated into it (See Table 4)

Countries 1 and 2 would seem to have higher injury rates than Country 3. Country 4 immediately presents as an outlier. For reference, all countries do similar work with similar risk profiles. (See Fig. 13)

Country 4 shows an injury performance that differs by orders of magnitude from that of the other countries. Two possible explanations for the considerable differences in rates are that Country 4 is substantially safer than its international counterparts or that reporting is substantially lower in that region (i.e., very low to non-existent). The additional severity metrics were neither necessary nor helpful in that assessment. Leadership suspected that reduced reporting from that world region was the likely cause, a narrative they already understood.

The more interesting narrative change surrounded Country 3. With 29 million hours and a TRIF of 1.99, there was sufficient data to suggest that reasonable reporting was occurring; the SAIF rate is of limited value, as it reflected the lower rate of events. The difference in average SAIF was the real indication that Country 3 is not being reasonably compared with Countries 1 and 2. The average event severity of 0.79 as compared to 0.41 and 0.42 in the other regions suggests that low- to moderate-severity events are not commonplace in the region. Whether this is a product of international cultural differences, medical practices, or reporting differences is not something this research can ascertain. What is clear is that the TRIF and SAIF rates are not based on the same event profiles, despite using the same definitions.

In discussions with the organization, there were no surprises about the rates for Country 4; the real epiphany was the severity reported by Country group 3. The group of countries that make up Country 3 were believed to be reporting in a manner similar to that of Countries 1 and 2. Safety leaders from the organization had guessed at differences in the nature of events but lacked a numerical basis for that inference. Demonstrating that the average event was 88% more severe for Country 3 provided clear numeric evidence that the TRIF numbers between regions were not inclusive of the same range of events. The general conclusion is that there was a material enough difference between countries' data inclusion to render a direct comparison of international TRIF numbers unreliable.

This finding would also then render a cumulative international score unreliable. Even if the biases for specific country groups remained consistent across periods, the data would not reflect expectations. The average event reporting differences indicate that cumulative data would need to be normalized by both hours and average events to accurately reflect an international business with culturally diverse event classifications.

3.3.3. Participant 3B – Overall story

Participant 3B was the Australian business for Participant 3. The general story is that the organization was on a relatively flat or slightly downward trend in TRIF until Jan 2024, after which its injury rate began to rise.

The first complete TRIF rate in the data set (with 12 months of data) is June 2023. Over this period, the organization has observed a general increase in injuries, from 11.05 to 12.75 (an increase of 15.4%). (See Figs. 14 and 15)

When SAIF and Average SAIF event scores are added, we see that the period from Jun 2023 through to December 2024 saw a 15.4% increase in injury rate, and a 16.7% reduction in the average severity of those events (0.48 down to 0.4).

In discussion with this business, there were initiatives in place at around that time to encourage positive response to events, in particular programs centered around Human and Organizational Performance (HOP) (Conklin, 2019). It was suggested that the increase in TRIF, coupled with a decrease in average severity, may reflect a more psychologically safe reporting culture.

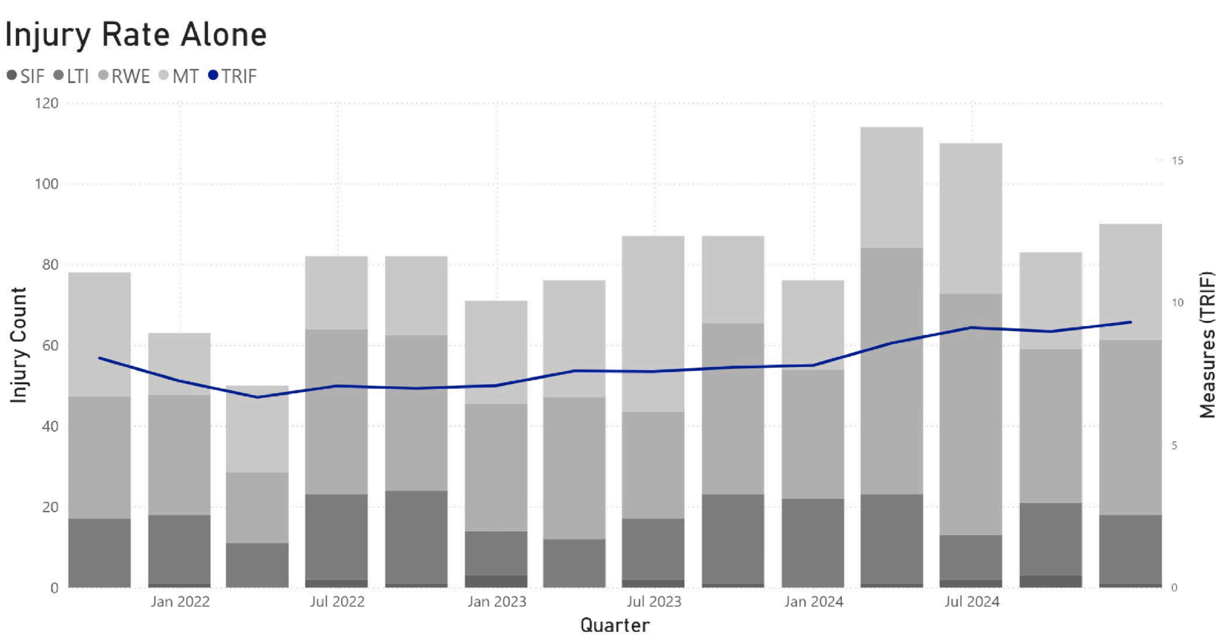

Fig. 11. Participant 3A – Multinational Full Company TRIF Alone (Quarterly).

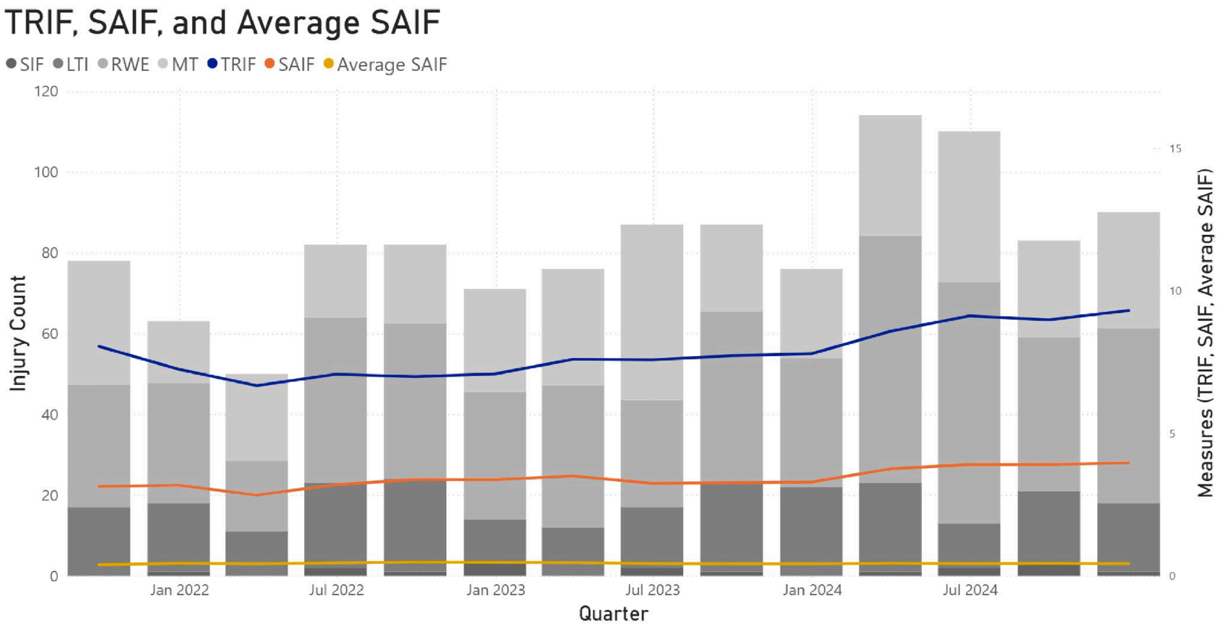

Fig. 12. Participant 3A – Multinational Full Company TRIF, SAIF and Average SAIF Rates.

Table 4
Participant 3A – Country TRIF, SAIF and Average SAIF – Q4 2024.

Country	TRIF	SAIF	Average SAIF	Total Hours
Country 1	11.54	4.72	0.41	90,395,921
Country 2	14.71	6.24	0.42	13,258,274
Country 3	1.99	1.57	0.79	29,454,456
Country 4	0.31	0.10	0.33	11,123,121
Total	**9.30**	**3.96**	**0.43**	**144,231,772**

3.3.4. Participant 3B – Mechanisms of injury

Participant 3A was unable to provide mechanism data for their international group but was able to do so for their Australian business. A similar pattern of failure mechanisms to those observed in previous case studies was observed. For Participant 3B, motor vehicle accidents were once again a low-frequency, but high-consequence event. (See Figs. 16 and 17)

This analysis focused specifically on the line-of-fire failure mechanism. These are events in which someone was injured while in the path of a moving object, equipment, or released energy. This analysis suggested that the event rate for this failure mechanism had increased over the past 19 months, with little change in the event profile. This finding was taken as strong evidence to support an upcoming safety program focused on controls for events of this nature. This analysis demonstrates the ability to track rate and relative severity for one particular failure mechanism.

4. Discussion and analysis

The design and implementation of injury measurement tools substantially altered the narratives of injury performance in both direction and complexity. When metrics were examined together, the narratives centered on understanding the relative relationships among injury measures rather than treating any single measure as the sole performance metric. The typical better-or-worse descriptions from stand-alone injury rates did not hold.

More narrative options were possible when measures were taken together and in context with one another. An increase in both the rate and severity of events could be interpreted as reduced performance; however, an increase in injury rate with a decrease in severity could indicate improved reporting culture; a decreasing injury rate with increasing severity could have the opposite interpretation. A conspicuous decline in both rate and severity raised questions about the source's trustworthiness when low-severity events were entirely absent.

The vast majority of identified injury rate trends were accompanied by an offsetting change in the relative severity. It would require considerably more data to mathematically prove an inverse correlation between injury rate and average severity. There is sufficient evidence to conclude that the majority of identified trends in injury rate were paired with offsetting changes in the average severity of those events. This has serious implications for the better-or-worse narrative of the injury rate. Most injury rate movements observed in this research were potentially driven by shifts in the inclusion or exclusion of lower-severity events.

For those looking at injury rates for their own organizations, the implication is this: When viewing any major trend in a stand-alone injury rate, it is likely that a change in the reporting of low-severity events is a better explanation of injury rate movement than a material shift in injury performance.

Measurement tools affected both the narrative's complexity and the corresponding discussions and responses. Paired measures prompted discussion of potential reasons for divergence between the measures, or why one measure held while another changed. Narratives about a paired set of measures were seldom discussed as better or worse. Rather, discussions moved towards culture and risk-based descriptions of the findings. The narrative explanations of injury performance changed to dynamic explanations of changing cultures.

The evaluation of failure mechanisms also benefited from a change in narrative. Low-volume, high-severity mechanisms require different controls from high-volume, low-severity failure mechanisms. The injury rate drives attention to high-volume events; a complete shift toward fatal risks drives safety programs to focus on high-severity events but may ignore moderate outcomes. Balancing measures of volume and severity enabled a more considered response in both cases.

Following this research, four key guidelines were developed that contributed to the overall ability to change the injury performance narrative. Those four guidelines are: Paired Analysis, Continuous Variability, a Simple Narrative, and Highlighting Differences.

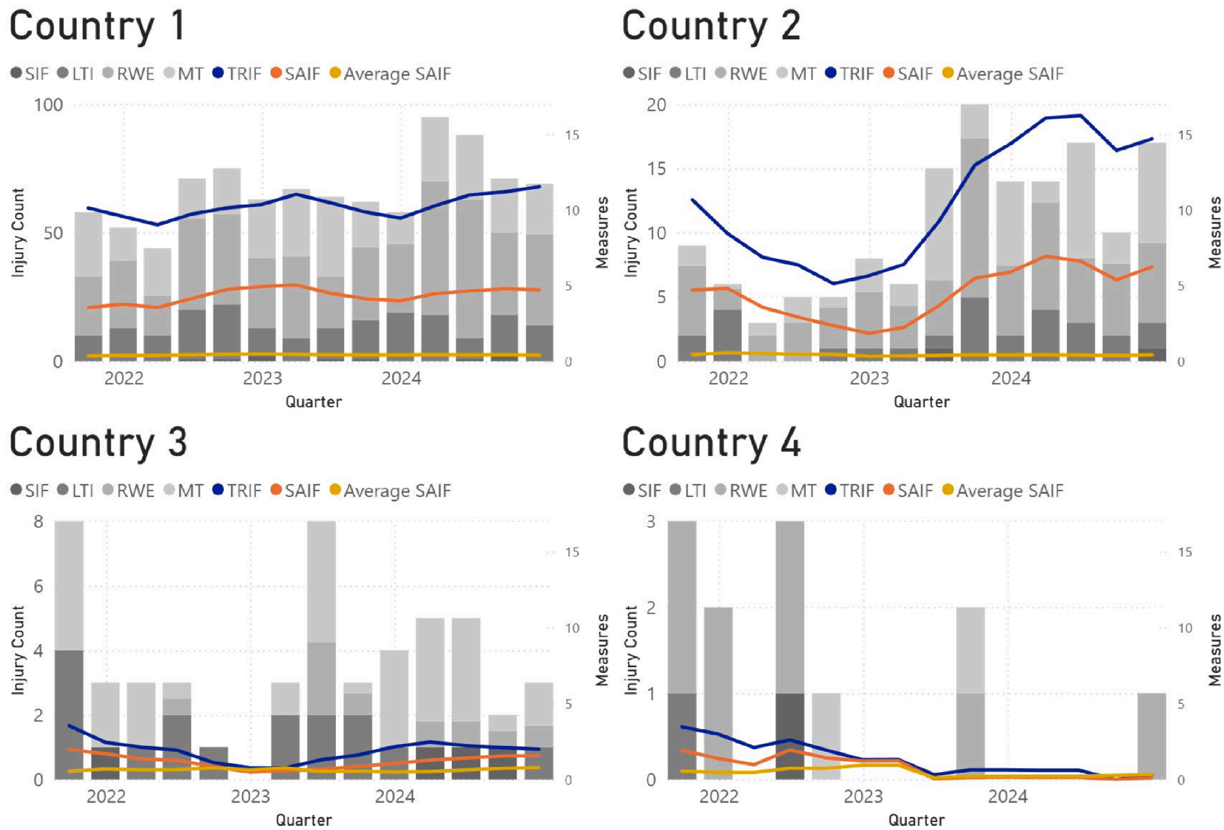

Fig. 13. Participant 3A – Country Comparison.

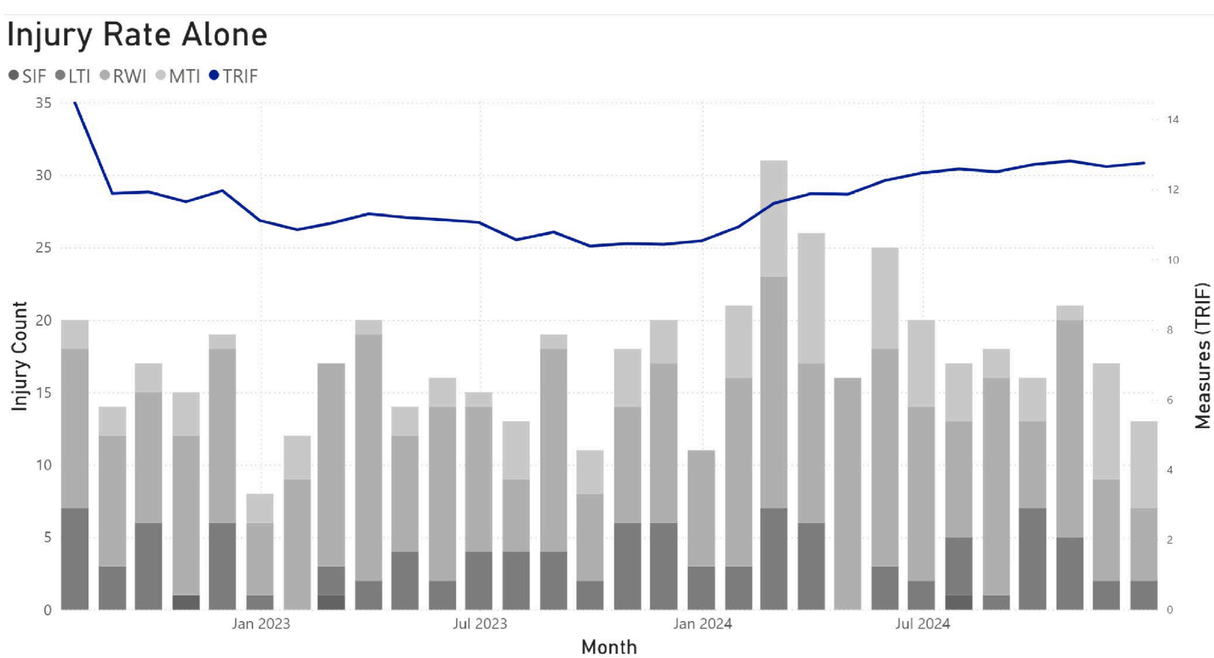

Fig. 14. Participant 3B – Australian Business TRIF Alone.

4.1. *Guideline 1: Pairing – injuries must be simultaneously assessed based on both rate and severity of outcome, with a weighted measure of the two for reference*

Through the three case studies, no single measure told the whole story of injury performance. Each of the three measures contributed to the injury performance narrative at different points. Different scenarios would feature one measure more prominently in the narrative, usually when one or more was moving while others were not. All three measures were, however, necessary to generate a complete narrative. Injury rate, average severity, and weighted frequency were necessary elements to a complete understanding of injury performance.

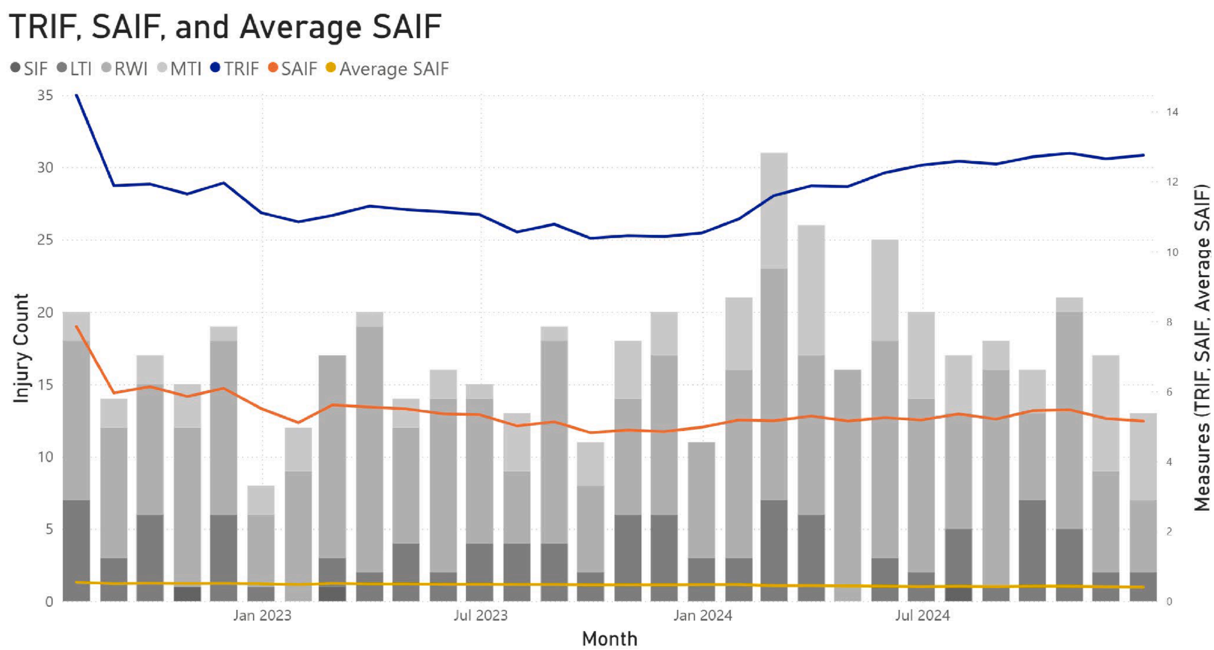

Fig. 15. Participant 3B – Australian Business TRIF, SAIF, and Average SAIF Rates.

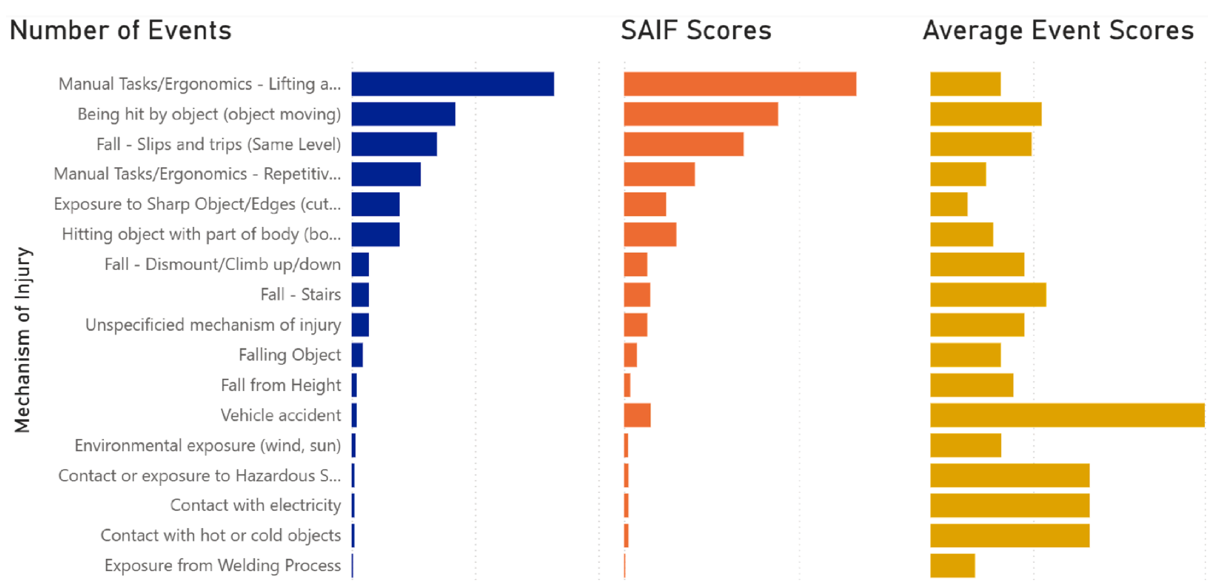

Fig. 16. Participant 3B – Frequency and Severity by Mechanism.

The production of single-measure charts was insufficient to have a reasonable understanding of the narrative behind injuries. Viewing the measures together was important in changing the narrative of any stand-alone rate. In practice, this meant plotting all three measures on the same chart so they could be viewed in the context of their relationships to other metrics.

Guideline: **Rate, Average Severity and a Weighted Average measure are not separable pieces of information. All three measures must be presented together to tell a complete narrative.**

4.2. Guideline 2: Continuous variability is important, particularly with lower-severity events

Participants responded favorably to the continuous variability within categories. Leaders indicated that the SAIF was sensitive to short-term LTIs, especially those with only one or two lost days. The response was positive, as the value clearly distinguished cases in which a single day changed an event's category. Participants suggested that changing the measurement tool would make a change in category less daunting than the current measures, which treat single days as equivalent to weeks or months of lost time.

The initial creators of the Injury Rate–Severity Rate pairing in 1916 recognized that underreporting would always be present and identified that a weighted frequency rate would mitigate its effects.

> *Accident reports are probably never absolutely complete, and, as a rule, the completeness of reporting is in direct proportion to the seriousness of injury. The more serious the injury the greater the likelihood of its being reported. Frequently the reporting of minor injuries is extremely incomplete…*
>
> *Inasmuch as the accuracy of frequency rates depends upon the completeness of accident reports, and as all accidents have the same*

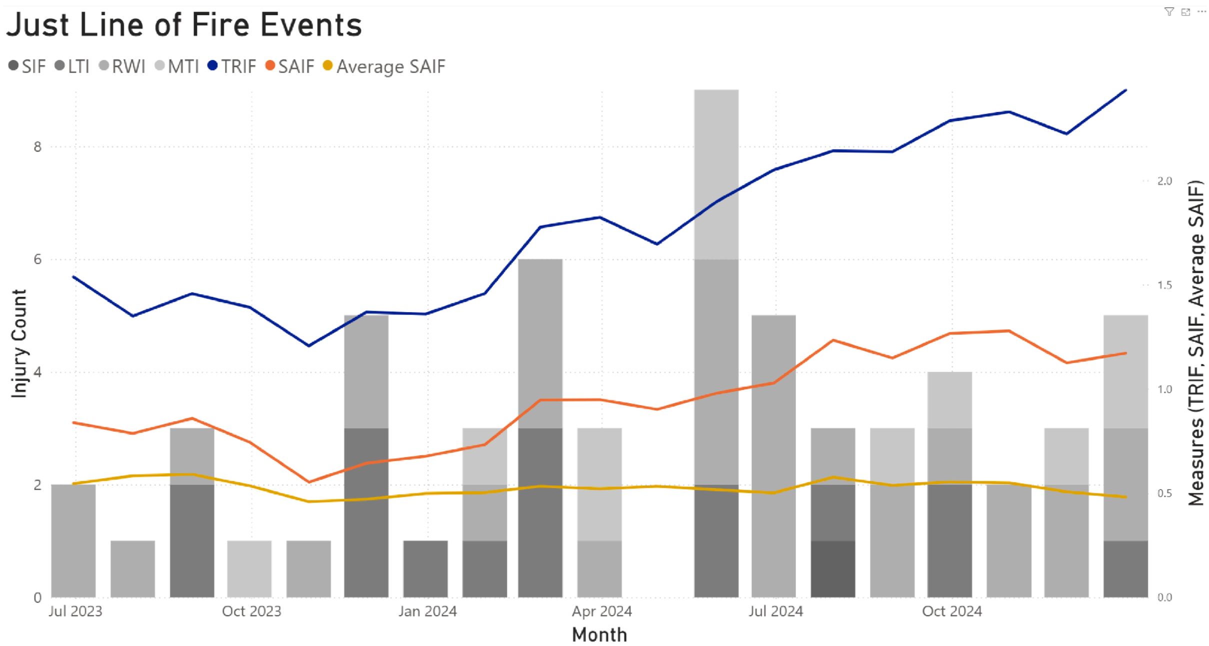

Fig. 17. Participant 3B – Line of Fire event trending.

> *weight, a failure to report any considerable number of minor accidents renders the rates obtained of very little value. Such is not the case with severity rates. Here the disabilities are weighted according to their importance and a large group of minor disabilities has comparatively little effect upon the derived severity rate.* Bureau of Labor Statistics, (1916)

.

Under-reporting is not an offense so much as it is a natural product of our measurement tools. As many as 80% of occupational injuries go unreported (De Silva et al., 2018; Geddert et al., 2021; Petitta et al., 2017; Probst et al., 2008; Probst & Estrada, 2010). In this research, the inclusion or exclusion of lower-severity events provided a mathematically sound explanation for many of the perceived trends in the standalone injury rate. A good set of injury measurement tools must acknowledge the reality of how measurement tools affect recording behavior.

Sharp category boundaries drive poor measurement behaviors. The inclusion of a sliding scale is important for making an injury assessment tool more effective in driving positive behavior. The objective is to create a safe reporting environment that minimizes the anxiety of involved parties associated with reporting (Dekker, 2017a).

Principle: **Under-Reporting is expected and is most prevalent in lower-severity injuries. This is moderated with a weighted frequency rate and the avoidance of categorical leaps, particularly in the value of lower-severity events.**

4.3. Guideline 3: A simple narrative

Clarity is an important characteristic of an effective measure (Erkal & Hallowell, 2023); particularly when the role of that measure is to tell a compelling story about injury performance. If a set of tools is to replace or repair the use of a standalone injury rate, those tools must convey a similarly clear message when assessing injuries for both rate and severity.

The SAIF rate, as defined in this work, is not the only means to report injuries in a paired and continuously variable manner. The SAIF rate model demonstrated benefits from the design feature of a benchmark injury, resulting in a simple communication narrative.

While pursuing this research, one organization applied the measure across their business. On numerous occasions, both leaders and safety professionals requested clarification on the SAIF rate definition. The response was consistently that TRIF denotes the number of people injured, and SAIF denotes the relative number with a two-week LTI. That explanation was immediately acceptable and actionable to just about everyone who asked. The simple narrative enabled users to engage with the measure, believing they understood what the data were telling them and how they could respond effectively.

Guideline: **The injury rate benefits greatly from being simply and intuitively understood; a measure of severity must be equally simple and intuitive to understand.**

4.4. Guideline 4: Severity measures must have an appropriate scale, that encompasses all injuries and highlights differences

It is important to recognize that the severity measure used in this research was not an accurate measure of severity, but rather a useful one for identifying differences between data sets. All events had some effect on the outcome of the measure, but none too much. If all events are approximately equal, then the measure has become the current recordable injury rate. If we scale the value of events so that high-severity events substantially outweigh lower-severity events, we have simply reproduced the SIF rate. In either case, the SAIF would tell us nothing new if it were not weighted in a manner that reasonably acknowledges all events and identifies category differences.

The original designers of the Injury Rate-Severity Rate pair in 1916 understood that what mattered was not the specific measure of severity but its ability to highlight differences.

> *The fact that the real significance of severity rates is in the measurement of relative hazards renders comparatively unimportant the character of the weighting scale used… It is desirable to have the scale used as accurate as possible but the fact that a completely accurate scale can not be devised does not impair the value of accident severity rating.* Bureau of Labor Statistics, (1916)

.

Learning from both the initial design of the severity rate in 1916 and its eventual failure, the fourth principle is that the severity scale must reasonably assess all events and indicate categorical differences.

It is important to note that a base-3 logarithm cannot truly honor the reality of suffering and loss experienced by someone who has undergone a life-altering event. It is indeed doubtful that quantifying the true reality of severity in terms of actual suffering and loss experienced would be either possible or helpful. The hope of this research is that, by creating usable measures of injury rates and severity, actual suffering and loss may be immeasurably reduced.

Principle: **A measure must assign a reasonable value to all associated events, but it is more important that a measure of severity be usable in highlighting differences than be an absolutely accurate measure of severity.**

5. Conclusions

The findings in this paper complement and extend the recent work of Pomeroy and Pilbeam (2025). Our work confirms that numerous meanings are generated from injury rates, and suggests that variation in how injury rates are presented has a direct impact on which meanings are generated and what responses occur as a result. Pomeroy and Pilbeam identified six uses of injury rates: a form of control, a means of motivation, a tool for self-promotion, a tool for measuring safety, a tool for measuring risk and noncompliance, and a sign of a leader's commitment. The ongoing question arising from our research is whether those six uses would yield the same meaning if a stand-alone injury rate as with alternative presentations involving severity. If moving away from stand-alone injury rates changes the narrative of injury performance, then it would logically follow that all six uses could be affected by that change. For example, leaders who demonstrated commitment with a low injury rate could instead do so by diligently reporting lower-severity events as reflected in their average severity. It would be an interesting question for future research to assess how each use case was affected by changing from a stand-alone injury rate to a more balanced set of injury measures.

This research defines four guidelines necessary to fundamentally change the narrative of a stand-alone injury rate. Our SAIF metric provides one way to meet these guidelines, but the use of SAIF in this study is as a demonstration only. Other measures may be equally capable of changing the narrative around injury performance if they adhere to the principles defined here for narrative-driven injury assessment.

Changing injury measurement tools affected more than just the direction of the narratives (positive or negative). It changed the complexity and character of the narratives as well. A significant limitation to this approach is that it relies on the interpretation of the person applying the measures. Measurement tools may enable more dynamic narratives to be generated, but they still impose an intellectual responsibility on those assessing the measures. Even with alternative measures, there were instances where one measure influenced the narrative more than others. Measures, narratives, and follow-up actions still require human professional and intellectual leadership to ensure that the decisions made have a meaningful impact on objective safety outcomes.

Changing the narrative of injury assessment is meant to be understood alongside a much broader definition of safety performance. Lagging indicators, such as these, may serve a purpose only if they are part of a larger approach to safety that encompasses leading indicators, measures of control, and measures of the capacity for things to go well. The narrative of safety, which extends beyond injuries, must address questions such as: have we controlled all major sources of energy release, and have we built capacity within our systems to adapt and deliver safe outcomes? Generating more dynamic injury narratives may help ensure that attention isn't disproportionately focused on high-volume, low-severity outcomes, or the reverse. It is, nonetheless, controls and capacity that will inevitably keep people safe. We will be led astray when we focus too much attention on one key measure, to the exclusion of others, including lagging injury measures rather than leading safety drivers.

This research demonstrates that injuries can be used to generate a simple narrative of differences in injury performance over time and between groups, or to identify differences in failure mechanisms. To be used effectively within a Safety Management System, a broader construct is needed that treats these rates as learning opportunities when our controls and capacity for positive outcomes fall short. The measure of controls, capacity, high-potential events, and interventions all contribute to understanding if we have effectively created the possibly unquantifiable condition we call safety.

Injury rates may yet make a significant contribution to the discussion of safety, but not on their own, and not as the definition of safety. The stand-alone injury rate, as a measure of "safety," is a tyrant driving under-reporting and poor real-world performance. The role of injury rates in a holistic safety measurement construct could be as follows: injury rates, paired with a severity rate and average severity, can provide insight into the real stars of safety: controls that stop things from going wrong, and the capacity to have things go right. For an organization to achieve safety, all those things must work together. To set any one measure as the new stand-alone definition of safety is to invite a new tyrant to enjoy the spot held so long by the stand-alone injury rate.

6. Statements of data, funding and generative AI

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

The data for this research have not been made available, as participation was to be anonymous. Any requests for data can be made to the corresponding author and must be assessed and approved by each research participant.

No AI tools were used in the preparation of this research or the writing of this paper. An express request was made that editors and reviewers also not use AI tools.

CRediT authorship contribution statement

Kevin Geddert: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sidney Dekker:** Writing – review & editing, Validation, Supervision, Conceptualization. **Drew Rae:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [I (Kevin Geddert) operate an independent data consultancy dealing primarily with financial and operational data analysis. Two of the anonymous participants have worked with the consultancy for purposes outside of this research.

My relationships with people at the organisations allowed me to encourage their participation in my research. My contractual work was not related to the research. I received no funding or consulting fees related to this research, and none of the participating organisations had any particular interest in influencing the results.

Nonetheless, I want to be transparent that I do unrelated consultative work in financial and operational data analysis, and two of the anonymous participants are organisations I previously worked with].

References

Anonymous. (1916). A new method of computing accident rates. *Monthly Review of the U. S. Bureau of Labor Statistics, 3*(1), 6–17. http://www.jstor.org/stable/41823025.

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

Coca-Cola. (2021). Coca-Cola 2020 Business & Environmental, Social and Governance Report. In (pp. 82).

Conklin, T. (2019). *Pre-accident investigations: an introduction to organizational safety*. CRC Press.

Consolazio, C. F. (1963). *Physiological measurements of metabolic functions in man. Blakiston Division.*

De Silva, N., Rathnayake, U., & Kulasekera, K. M. U. B. (2018). Under-reporting of construction accidents in Sri Lanka. *Journal of Engineering, Design and Technology, 16*(6), 850–868. https://doi.org/10.1108/JEDT-07-2017-0069

Dekker, S. (2017a). *Just culture : Restoring trust and accountability in your organization* (Third edition. ed.). CRC Press, Taylor & Francis Group.

Dekker, S. (2017b). Zero vision: Enlightenment and new religion. *Policy and Practice in Health and Safety, 15*(2), 101–107. https://doi.org/10.1080/14773996.2017.1314070

Eagan, C. (1964). The effect of air movement on atmospheric cooling power. *Review of research on military problems in cold regions. Ft. Wainwright, Alaska: Arctic Aeromedical Laboratory TDR-64-28*, 147-156.

Erkal, E. D. O., & Hallowell, M. (2023). Moving beyond TRIR: Measuring & monitoring safety performance with high-energy control assessments. *Professional Safety, 68*(05), 26–35. https://onepetro.org/PS/article-abstract/68/05/26/519591/.

Fu, G., Xie, X., Jia, Q., Li, Z., Chen, P., & Ge, Y. (2020). The development history of accident causation models in the past 100 years: 24Model, a more modern accident causation model. *Process Safety and Environmental Protection, 134*, 47–82. https://doi.org/10.1016/j.psep.2019.11.027

Geddert, K. (2025a). Analysing safety data for an entire industry group – a case study from Queensland Coal-Seam Gas. *Australian Energy Producers Journal, 65*. https://doi.org/10.1071/EP24485

Geddert, K. (2025b). Meeting the target, missing the point. In S. Dekker (Ed.), *Safety Theater: How the desire for perfrection drives compliance clutter, inauthenticity and accidents - and what to do about it*. Routledge.

Geddert, K., Dekker, S., & Rae, A. (2021). How does selective reporting distort understanding of workplace injuries? *Safety, 7*(3), 58. https://www.mdpi.com/2313-576X/7/3/58.

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.

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

Hollnagel E, W. R., Braithwaite J. Middelfart,. (2015). From Safety-I to Safety-II: A White Paper. https://psnet.ahrq.gov/issue/safety-i-safety-ii-white-paper.

Hopkins, A. (2009). Thinking about process safety indicators. *Safety Science, 47*(4), 460–465. https://doi.org/10.1016/j.ssci.2007.12.006

Orvik, C. P., Albrechtsen, E., Kongsvik, T., Holen, S. M., Bertheussen, L. E., & Moltubakk, S. T. (2026). What is safety performance? A systematic review of conceptualizations in the construction safety research. *Safety Science, 193*, Article 107025. https://doi.org/10.1016/j.ssci.2025.107025

Osczevski, R., & Bluestein, M. (2005). The new wind chill equivalent temperature chart. *Bulletin of the American Meteorological Society, 86*(10), 1453–1458. http://www.jstor.org.libraryproxy.griffith.edu.au/stable/26221366.

Osczevski, R. J. (1995). The basis of wind chill. *Arctic, 48*(4), 372–382. http://www.jstor.org.libraryproxy.griffith.edu.au/stable/40511941.

Petitta, L., Probst, T. M., & Barbaranelli, C. (2017). Safety culture, moral disengagement, and accident underreporting. *Journal of Business Ethics, 141*(3), 489–504. https://doi.org/10.1007/s10551-015-2694-1

Pomeroy, J., & Pilbeam, C. (2025). Signs of safety: An investigation of how OHS professionals interpret injury metrics. *Journal of safety Research, 95*, 87–100. https://doi.org/10.1016/j.jsr.2025.09.010

Probst, T. M., Brubaker, T. L., & Barsotti, A. (2008). Organizational injury rate underreporting: The moderating effect of organizational safety climate. *Journal of Applied Psychology, 93*, 1147–1154. https://doi.org/10.1037/0021-9010.93.5.1147

Probst, T. M., & Estrada, A. X. (2010). Accident under-reporting among employees: Testing the moderating influence of psychological safety climate and supervisor enforcement of safety practices. *Accident Analysis & Prevention, 42*(5), 1438–1444. https://doi.org/10.1016/j.aap.2009.06.027

Probst, T. M., & Graso, M. (2013). Pressure to produce=pressure to reduce accident reporting? *Accident Analysis & Prevention, 59*, 580–587. https://doi.org/10.1016/j.aap.2013.07.020

Provan, D. J., Woods, D. D., Dekker, S. W. A., & Rae, A. J. (2020). Safety II professionals: How resilience engineering can transform safety practice. *Reliability Engineering & System Safety, 195*, Article 106740. https://doi.org/10.1016/j.ress.2019.106740

Raheemy, Y., Sherratt, F., & Hallowell, M. R. (2025). What is safety? contemporary definitions and interpretations across North America. *Safety Science, 185*, Article 106798. https://doi.org/10.1016/j.ssci.2025.106798

Shell. (2022). Strategic Report 2021. In.

Siple, P. A., & Passel, C. F. (1945). Measurements of dry atmospheric cooling in subfreezing temperatures. *Proceedings of the American Philosophical Society, 89*(1), 177–199. http://www.jstor.org.libraryproxy.griffith.edu.au/stable/985324.

Smith, R., & Jones, M. (2014). A hurt-based approach to safety. *IPTC 2014: International Petroleum Technology Conference*.

Kevin Geddert is a researcher and data analytics consultant. Kevin's experience in measurement systems spans diverse areas such as Safety, Finance, Human Resources, Maintenance systems, Operations reporting, and others. Kevin runs and independent consultancy on data and analytics, is a regular presenter with the industry group Safer Together, and works to collate safety data from across Australia relating to the Coal Seam Gas industry. He started working with Griffith University to help further research into the challenges in safety reporting and opportunities to research the measurement of injuries.

Sidney Dekker A professor, mediator, chaplain, author, pilot, and powerful speaker, Sidney's won worldwide acclaim for pushing the boundaries in human factors and safety. Born near Amsterdam as the middle child of a Government Director and one of the first female ministers in the Dutch Reformed Church, Sidney grew up drawing inspiration from his mum's sermons and pastoral care-giving, and his dad's pragmatic and compassionate human-centred leadership.Sidney is in the top 1% most influential scientists since Isaac Newton (Standford ranking), his work has over 20,000 citations. He has published 21 books and given more than 500 talks or keynote addresses globally across a range of industries.

Dr Drew Rae is an Associate Professor in the Safety Science Innovation Lab at Griffith University, where he teaches courses on research methods and safety engineering, and manages the lab's research program. Drew's own research uses a mix of ethnography, field experiments and theory-building to investigate organisational safety practices. He is particularly interested in understanding the myths, rituals and bad habits that surround the work of managers and safety practitioners, and how this work influences front-line operations. Drew co-hosts the Safety of Work podcast and is on the editorial board of the journal Safety Science.