International benchmarking of railroad safety data systems and performance – a cross-continental case study

Chen-Yu Lin\textsuperscript{a,*}, Marcelo Blumenfeld\textsuperscript{b}, Theodore Gerstein\textsuperscript{c}, Christopher P.L. Barkan\textsuperscript{c}, Anson Jack\textsuperscript{b}, Usman T. Abdurrahman\textsuperscript{b}

\textsuperscript{a} Department of Transportation and Logistics Management, National Yang Ming Chiao Tung University, Assembly Building 1, No.1001, Daxue Rd., East Dist., Hsinchu City, 30010, Taiwan
\textsuperscript{b} Birmingham Centre for Railway Research and Education – BCRRE College of Engineering and Physical Sciences, University of Birmingham, 12 Edgbaston Park Rd, Birmingham, B15 2TT, United Kingdom
\textsuperscript{c} Rail Transportation and Engineering Center – RailTEC, Department of Civil and Environmental Engineering University of Illinois at Urbana-Champaign, 8118 Newmark Engineering Laboratory, 205 North Mathews Avenue, Urbana, IL, 61801, USA

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Rail safety
Benchmarking
Train accident
Data system
Risk management

\textbf{ABSTRACT}

Rail safety is a universal goal that every railroad system pursues. Comparing rail safety performance allows the identification of relative system strengths and weaknesses and potential adaptation of risk mitigation strategies from one railroad system to another. Achieving this requires comparable data from the railroad systems to be analyzed that are available and with high resolution so that fair comparisons can be established. This paper presented an international benchmarking framework for railroad safety-related data system and safety performance. A novel and standardized methodology was developed to collect railroad safety-related data sources among different countries and compare their data completeness and resolution. Six countries with high data availability and transparency were selected to demonstrate the benchmarking framework. High-level rail safety performance measures were derived and compared among these countries. The results showed that there are inconsistencies in the resolution of different types of rail safety data among the six countries. The countries that had the lowest and highest overall accident rate, grade crossing incident rate, and other safety performance metrics were identified. This research provided valuable insights into how railroad operators can improve their railroad safety-related data system and safety performance by benchmarking with other railroad operators to make the most efficient use of risk mitigation resources. The study also highlighted the importance of providing publicly available railroad safety-related data for the mutual benefits of overall safety of railroad systems around the world.

1. Introduction

1.1. Rail safety – a universal goal

Safety should be the first priority of any railroad system. A safe and reliable rail transportation environment in a country not only benefits individual train operators but also the nation’s rail transport development. The focus to improve rail safety is worldwide and

* Corresponding author.

\textit{E-mail addresses:} chenyulin@nycu.edu.tw (C.-Y. Lin), M.Blumenfeld@bham.ac.uk (M. Blumenfeld), tsg4@illinois.edu (T. Gerstein), cbarkan@illinois.edu (C.P.L. Barkan), a.c.r.jack@bham.ac.uk (A. Jack), usmantee@gmail.com (U.T. Abdurrahman).
every country devotes tremendous resources to reduce railroad accident and incident risks. Major safety improvement programs have been undertaken and advanced technologies implemented to address the most important rail safety challenges. For example, the United Kingdom initiated Leading Health and Safety on Britain’s Railway, a cross-industry strategic group with the focus of identifying and addressing the most significant rail safety problems in the country (Rail Safety and Standards Board (RSSB), 2021). In North America, the United States is deploying Positive Train Control (PTC) systems to reduce the risks of train collisions, overspeed derailments, accidents caused by misaligned switches, and maintenance-of-way (MoW) incidents (Zhang et al., 2018a). Different countries’ rail systems may vary in the characteristics of their infrastructure, rolling stock, signaling, train control systems, operations, and organizational structures, resulting in differing safety issues. For example, on a rail system with frequent, heavy-haul freight train operation, infrastructure problems may be more important (Wang et al., 2020; Khajehei et al., 2021). Note that the word “system” here refers to the integrated components and elements that enable the operation of trains, including infrastructure, rolling stock, signal and train control, people, and other necessary hardware and software. The term “railroad system” in this study refers to a country’s overall rail network and its train operation.

Although railroad system and operating features may vary, the fundamental physics of train accident mechanisms remain the same. In other words, if a factor, or combination of factors, can cause an accident, the event is likely to occur regardless of other unrelated features. For example, a broken rail derailment is likely to occur wherever a train runs through a defective rail section whose structural integrity is compromised, regardless of other system characteristics. Therefore, the efforts that one system invests to mitigate that risk may be adapted by another with appropriate adjustments, while acknowledging that rail systems consist of interacting elements and subsystems. On the other hand, if a railroad changes or adds system features, new risk mitigation measures may be required to maintain or improve the level of safety due to these changes. In such cases an operator can reference existing systems with similar features and proactively adapt their approach to mitigate the risks. For example, if a freight-dominant railroad system introduces passenger train services, risk mitigation measures specifically for the operation of passenger trains and shared passenger and freight corridors can be learned from existing passenger-train dominant and shared corridor systems that have already developed risk mitigation measures. Learning from other systems who have mature, well-established strategies can prevent the same types of risk thereby skipping some “growing pains”. That said, adaptation among different systems can be challenging because it requires motivation to change, openness, and willingness to learn from others (Ki, 2021), based on evidence of a more effective rail data system or safety performance.

1.2. Benchmarking rail safety data and performance

A crucial step to facilitate the aforementioned adaptation is benchmarking, which is the process of identifying the highest standard of excellence for products, systems, operations, or services, and then making the improvements necessary to reach those standards – commonly known as “best practices” (Bhutta and Huq, 1999). A general benchmarking framework consists of five phases: Plan - determine what to benchmark and benchmark methodology, Find - identify benchmarking partners, Collect – gather necessary data and information for benchmarking, Analyze – identify gaps in performance of benchmarking topics among different partners, and Improve – develop suggestions and plans for improvements (Lai et al., 2011). These phases are commonly illustrated in a “benchmarking wheel” for its iterative nature (Fig. 1).

Benchmarking rail safety-related data and safety performance among different railroad systems can elucidate relative strengths and weaknesses and suitable risk mitigation measures implemented. If the necessary comparisons are to be informed by analysis of different railroads’ safety data systems, then these systems must be reasonably consistent and have high resolution. While most railroad operators collect at least some safety-related data, such data are often not publicly available, and the level of those data are
often not consistent. Although collaboration between consenting railroad operators to share, analyze, and compare railroad safety performance may be possible, public access offers greater transparency. It enables other parties and researchers to conduct independent analyses, thereby encouraging innovative approaches and potentially lending greater credibility to the results. Such data availability and transparency allow objective benchmarking of rail safety performance among different countries, sharing of best practices, and encourages development and validation of more powerful, generalized railroad accident prediction and simulation models as well as implementation of risk mitigation strategies (Blumenfeld et al., 2023).

1.3. Research objectives

This paper presents an international benchmarking framework for railroad safety-related data system and safety performance. A novel and standardized methodology was developed to collect railroad safety-related data sources among different operators (countries) and derive safety statistics. Six countries in three different continents with high data availability and transparency were selected to demonstrate the benchmarking framework. Data availability is defined as the availability of safety-related data of a railroad system, and data transparency is defined as how much of the data are publicly available (Blumenfeld et al., 2023). High data availability and transparency indicates more comprehensive publication of key rail safety performance indicators such as fatality and injury rates, train accident rates, level crossing incident statistics, and so on (Blumenfeld et al., 2023).

Two benchmarking topics are focused on in this study.

1) Rail safety-related databases
2) Performance indicators available for evaluating rail safety

Data completeness and resolution among these countries were examined, and high-level rail safety statistics derived and compared. This research aimed to provide insights into how railroad operators can improve safety-related data system and safety performance by benchmarking with other railroad operators, and how the benchmarking framework helps to identify data system strengths and weaknesses and highlight the importance of providing publicly available and high-resolution data. The best practices of data collection for deriving a comprehensive set of rail safety performance are provided based on the benchmarking result.

1.4. Literature review

In this section, previous studies that benchmarked railroad safety data systems and safety performances were examined. While general benchmarking methods have been well established, there have been few studies benchmarking the quality and resolution of rail safety databases. Blumenfeld et al., 2023 developed a two-layer framework for benchmarking railway safety at the international level, creating an index system to evaluate rail safety data availability, transparency and publishing safety performance indicators. The authors are unaware of any other comparative analyses that focused on rail safety data systems and the consequent implications for comparison of safety performance. There are some organizations formed by a group of rail operators for benchmarking the performance of their systems such as the International Suburban Rail Benchmarking Group (ISBeRG), but the scope of the benchmarking is limited to participating rail systems, and the data collected and shared are confidential and focused on train delay and service reliability (Karathodorou and Condry, 2016; Piner and Condry, 2017).

There are some studies comparing rail safety performance among different railroad systems, and many of them focused on benchmarking European rail networks. This is likely because safety-related data in these countries are well organized by standardized, centralized data systems such as those maintained by the European Union Agency for Railways (EUAR). This provides homogeneous data for comparative analysis. For instance, Evans (2011) conducted a statistical analysis of fatal train accidents for European countries and compared their accident rate, severity, and causes. Sangiorgio et al. (2020) implemented a hybrid Analytic Hierarchy Process (AHP) – optimization framework to compare rail safety performance among European rail networks. Outside of the European continent, studies used existing rail safety databases to conduct quantitative risk analyses to compare safety performance. Kyriakidis et al. (2012) conducted a comparative analysis of railway accident precursors for multiple metro systems around the world. Rungskunroch et al. (2021) used Bayesian networks to benchmark the level of safety for high-speed rail systems in five countries accounting for uncertainty and lack of accident data. These studies focused on developing and comparing rail safety performance using the existing available data, but they did not delve into the data systems themselves nor consider how improvements in these systems might lead to more accurate and comparable safety performance measures. The scope of benchmarking of existing studies were bounded by the availability of rail safety data and the level of detail of these data.

Previous work benchmarking the safety performance of other transportation modes and multi-modal transport in different countries and systems was also reviewed. Barnett (2020) conducted a statistical analysis comparing aviation safety in different countries and regions of the world. Savage (2013) and Liu and Moini (2015) investigated and compared the fatality rates of different transport modes in the United States. Hermans et al. (2009) and Shen et al. (2020) benchmarked highway safety among multiple European countries using data envelopment analysis (DEA) methodology. Shen et al. (2015) reviewed and summarized various roadway safety indices developed for international benchmarking. MacLean et al. (Maclean et al., 2016) developed a stochastic-based method to evaluate aviation safety based on existing, available failure data. Similar to rail safety benchmarking, these studies focused on benchmarking safety based on existing data available and did not benchmark safety databases themselves. There has been little research that analyzed and benchmarked the structure and resolution of safety-related databases to identify the best practices of safety data collection and organization.
The result of the literature review is summarized below.

- The scope, level of detail, and accuracy of rail safety performance analysis and comparison depend on the availability, resolution, and homogeneity of safety-related data
- Previous research conducted risk analyses for rail systems based on existing and available data only and did not explore the quality of databases themselves
- The relationship between the availability and resolution of rail safety data and corresponding risk analyses and methods for evaluating safety performance is not well understood
- There is a need for a benchmarking of rail safety data structure and resolution to assess the strengths and weaknesses of individual rail safety data systems and provide suggestions for best practices of data collection to facilitate the most comprehensive and accurate evaluation of rail safety performance

The research presented in this paper addresses the need to benchmark rail safety databases and identify gaps in rail safety data collection practices and usage of these data. Among the contributions of this paper is a systematic examination and comparison of different nation’s rail safety data systems in a manner not previously been done. The results will provide insight to rail operators in how to construct databases more useful for conducting risk analyses and informing rail safety improvements.

2. Methodology
2.1. Benchmarking framework

The research framework of this study consisted of five major steps, corresponding to the five phases in a benchmarking wheel customized with rail-safety-specific elements (Fig. 2). First, the benchmarking topics were determined to be rail safety data systems and available safety performance indicators. Second, a set of countries whose rail safety-related data had high data availability and transparency were identified as benchmarking partners (Blumenfeld et al., 2023). Third, rail safety data of these countries were collected and organized into traffic exposure data and accident/incident data. Traffic exposure data provided insights into railroad traffic composition and allowed normalized railroad operational risk evaluation and comparison among different countries. Accident/incident data included the frequency of various types of train accidents/incidents and accident consequences such as passenger injuries and fatalities. In the fourth step, a thorough examination was performed to each country’s railroad safety data systems to identify similarities and differences in train accident/incident reporting and database construction processes, because these affect the resulting risk evaluation and interpretation. Then, based on the common data resolution, several railroad safety performance measures were developed and compared. Finally, based on each country’s railroad safety performance, consideration was given to potential improvements in accident/incident data collection and database organization, safer operating rules and practices, implementing alternative risk mitigation strategies, and adapting foreign technologies. These discussions focused on how railroad systems with better safety performance can provide useful information to systems that need improvements. It is possible that one system performs better than another in one safety measure but less well in another. In this respect the two systems can provide mutually beneficial information to improve safety.

Six countries were selected for benchmarking: the United States (US), the United Kingdom (UK), Germany (DE), France (FR), Sweden (SE), and Australia (AU), because they have reliable publicly available data, meaning that the data are issued by either governmental agencies or an officially recognized institute such as EuroStat, a Directorate-General of the European Commission. The selection also represented a variety of railroad operational environments: passenger versus freight centric system, predominantly private versus public ownership, and different geographic areas. These sources were also selected for their high data resolution. Data
resolution refers to the level of detail that rail safety data are broken down. For instance, while assessing casualty data related to train accidents, some systems provide a breakdown of casualties as minor injuries, major injuries, and fatalities, while others only provide a single figure for total casualties. Therefore, the former data were deemed to have higher data resolution than the latter in terms of casualties. In this study, there are 28 indicators (data types) in three data categories (Traffic data, accident/incident data, and consequence data) for measuring data resolution (Table 1), and the countries selected all have at least 60% (17 or more) indicators available, meaning that they have sufficient data resolution to warrant more in-depth analyses. Other countries either do not have the same level of data resolution, or part or all of their data are confidential or not recorded in English. The next subsection introduced all sources of public data identified from 2015 to 2019 for each country’s railroad system. The timeframe was selected because most of the data prior to 2015 were either not publicly available or incomplete for these countries. The types of data examined included traffic, accident/incident frequency and consequences, and near misses and accident precursors. Accident precursors refer to events or activities that create dangerous situations that may lead to a train accident. For example, an engineer (train driver) overrunning a stop signal is deemed as a precursor, because this event puts the train in a hazardous situation in which a train collision could ensue. Whether or not this situation, usually referred to as a signal passed at danger, or SPAD, results in a train collision, a near miss (the train collision almost happened but did not), or a safe end state, the occurrences of these precursors represent potential risks in railroad operational safety.

2.2. Data sources

2.2.1. United States (US)

Railroad traffic data were collected from the US Department of Transportation (DOT), Federal Railroad Administration (FRA), with supplemental information from the US DOT Bureau of Transportation Statistics (BTS), and the Surface Transportation Board (STB). These included freight and passenger train miles, freight ton-miles, and passenger miles accrued on the “general railroad system of transportation (the general system)” that is under the FRA’s jurisdiction (49 CFR Appendix A to Part 209, 2011). Light rail systems, transit systems, and isolated excursion rail lines were excluded.

The US accident/incident data were collected from multiple FRA accident databases. The Rail Equipment Accident and Incident (REA) database records railroad accidents that resulted in the total rail infrastructure and equipment damage exceeding a monetary threshold ($11,500 in 2023) which is periodically adjusted for inflation (Federal Railroad Administration (FRA), 2023). These accidents are called “reportable accidents”. Highway-rail Accident/Incident (HRA) database records all grade crossing incidents regardless of the total monetary damage. Railroad Injury and Illness (RII) database records all reportable human injuries and fatalities related to railroad operations. These three FRA databases are complementary to one another and in combination form a comprehensive safety

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Resolution Indicator</th>
<th>Type of Data</th>
<th>Resolution Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Data</td>
<td>Total Train-Miles, Freight Train-Miles, Passenger Train-Miles, Non-Mainline Train-Miles, Freight Ton-Miles, Freight Car-Miles</td>
<td>Consequence Data</td>
<td>Total Casualty Data, Passenger Casualty Data, Workforce Casualty Data, Non-Railroad Casualty Data, Number of Railcars Derailed, Hazardous Materials Release Consequences, Distinction Between Passenger and Freight Train Casualties, Distinction Between Mainline and Non-Mainline Casualties, Distinction Between Different Levels of Injuries</td>
</tr>
<tr>
<td>Accident/Incident Data</td>
<td>Passenger Miles, Reporting Threshold, Accident Type Breakdown: Derailment, Collision, Highway-Rail Grade Crossing, Train Fire, Obstructions (Striking Objects in Non-grade-crossing Areas), Hazardous Materials Releases, Distinction Between Passenger and Freight Train Accidents, Distinction Between Mainline and Non-Mainline Train Accidents, Trespasser Incidents, Railroad Suicide Data, Precursor/Near Miss Data</td>
<td>Notes:</td>
<td>1. The data resolution for a country is measured by the number of indicators that the country’s rail safety data system records. 2. There are 7 indicators in traffic data category, 12 indicators in accident/incident data category, and 9 indicators in consequence data category, forming a total of 28 indicators. 3. Recording of each type of accident/incident counts as one indicator, i.e., if a country records derailments and collisions separately, it gains 2 data resolution points.</td>
</tr>
</tbody>
</table>
database. The REA database collects all types of reportable accidents and therefore is the one that is most referred to and used for risk analyses. For each accident recorded in REA, there are over 140 data fields that document various characteristics of the accident, including track information, train consist, operational characteristics, casualties, monetary damage, and so on. The other two FRA accident databases provide more detailed documentation of information for a specific aspect or scope of all accidents. For example, HRA database only records grade crossing (level crossing) accidents but provides more detail on the grade crossing than the REA does (e.g., the angle between roadway and railroad track). Another difference between the HRA and REA is that the HRA also records all grade crossing accidents while the REA only records reportable grade crossing accidents. The RII database only records information related to human casualties but in more detail than the REA does (e.g., gender and age). Similar to the traffic data, the accident/incident data collected, and corresponding safety statistics produced, represented the rail safety performance of the general system in the US. No near miss and precursor data are available in the FRA databases as there is no regulation that required the mandatory reporting of such data.

The railroad accidents and incidents considered in the evaluation of US rail safety performance in this study included the REA-reportable derailments, collisions, highway-rail grade crossing incidents, obstructions, train fires, and other miscellaneous REA-reportable accidents on the general system. HRA database was used with the REA database to produce the comprehensive grade crossing safety statistics. Note that only reportable grade crossing accidents are considered in the accident statistics, meaning that non-reportable grade crossing accidents in the HRA are not included, for consistency and fair comparison when compared to other types of accidents. Non-train-related railroad incidents (e.g., trips and falls in train stations) were excluded. Note that trips and falls incidents generally refer to people tripping and falling at non-track areas within rail facilities (mostly inside train stations), e.g., escalator, elevator, stairs, wet floors, etc. If people fall and intrude onto railroad right-of-way, these incidents are generally considered as rail trespass events. Rail trespass incidents also include people deliberately foul railroad right-of-way at non-grade-crossing areas.

### 2.2.2. United Kingdom (UK)

The Office of Rail and Road (ORR) in UK collects rail safety data for the conventional railway system (“the mainline rail network”), London Underground network, and other non-mainline rail networks (Office of Rail and Road (ORR), 1230). The mainline rail network resembles the scope of the general railroad in the US. There is no reporting threshold for an accident or incident to the ORR, so the collection of accident data is more comprehensive, but this makes it less comparable to the US data. The main source for reporting accidents and incidents on the mainline rail network for the ORR is the Rail Safety and Standards Board (RSSB). Unlike the US, the RSSB records near miss and precursor data in cases where an event could potentially lead to an accident or incident.

A subset of the UK rail safety data from the ORR is reported to the EUAR, including significant accidents. A significant accident is defined by the EUAR as “Any accident involving at least one rail vehicle in motion, resulting in at least one killed or seriously injured person, or in significant damage exceeding £150,000, or extensive disruptions to traffic” (European Union Agency for Railways (EUAR), 2020). Any occurrence recognized as a significant accident is reported to the European Railway Accident Information Links (ERAIL) maintained by the EUAR for the 29 reporting entities including European Union member states, UK, Norway, and Switzerland (excluding Malta and Cyprus that do not have rail networks) (European Union Agency for Railways (EUAR), 2020). Accident severity data, including fatalities, injuries, and total damage are also reported to the ERAIL database. Rail traffic data for the UK were collected from the ORR, EUAR, and Eurostat.

The railroad accidents and incidents considered in the evaluation of UK rail safety performance in this study included EUAR-reportable derailments, collisions, highway-rail grade crossing incidents, obstructions, train fires, explosions, and other miscellaneous EUAR-reportable accidents on the mainline rail network. The EUAR instead of ORR data were used to allow fair comparison with other European countries. Non-train-related railroad incidents were excluded. Although the reporting thresholds between EUAR and FRA are different, both imply certain criteria for train accidents to be considered occurrences that should be prioritized in risk assessment and will help improve railroad system safety most effectively. Near miss and precursor data were also collected, not for direct comparison, but for discussions of rail safety data completeness and resolution. Traffic and accident data for light rail systems, transit systems, and isolated excursion rail lines were excluded for consistent comparison with other countries.

### 2.2.3. Germany (DE), France (FR), and Sweden (SE)

Railroad traffic data for France, Germany, and Sweden were referenced from Eurostat. The data collected consisted of freight and passenger train miles, freight ton-miles, and passenger miles. Traffic from light rail, transit, and excursion trains were excluded. Accident and incident occurrences as well as fatality and injury data were obtained from the EURA databases. Although EURA also collects near miss and precursor data, they are not publicly available. There is no publicly available source for rail safety data in English in individual countries.

Similar to UK, EURA-reportable derailments, collisions, highway-rail grade crossing incidents, obstructions, train fires, and other miscellaneous EURA-reportable accidents for these three countries were considered in this study. However, of these accidents there was no distinction between accidents occurring on passenger or freight trains. Therefore, safety measures developed for these countries were aggregated and not broken down by train type.

### 2.2.4. Australia (AU)

Australian passenger and freight-train-mile traffic data were provided by the Office of the National Rail Safety Regulator (ONRSR) (Office of the National Rail, 2021), and their passenger mile and freight ton-mile data were sourced from the World Bank (The World Bank, 2021). ONRSR was established in 2012 to provide a consistent risk-based application and regulatory oversight of rail safety in all Australian states and territories. Since states joined ONRSR progressively (the last being Queensland in 2017), train mile data prior to
2017 did not include all the states. Traffic data from light rail, transit, tourist, and heritage rail lines were excluded for consistent comparison with other countries.

Railroad accident/incident, injury, and fatality data were also provided by the ONRSR. An ONRSR-reportable occurrence is defined as “An incident associated with railway operations that has, or could have, caused significant property damage, serious injury or death.” (Office of the National Rail, 1951). There is no reporting threshold for an accident or incident to the ONRSR, so the collection of accident data is potentially more comprehensive, but this makes it less comparable to the data that have a specified reporting threshold. In this study, the ONRSR-reportable derailments, collisions, highway-rail grade crossing incidents, obstructions, train fires, and other miscellaneous accidents were considered. Similar to the European nations, data on the type of train these accidents occurred on is limited. Near miss and precursor data were also collected in this study, again not for direct comparison, but for consideration of rail safety data completeness and resolution.

2.2.5. Data source summary

The publicly available rail safety-related data sources for all six countries from 2015 to 2019 were summarized and compared (Table 2). Data prior to 2015 for some countries was either unavailable or unreliable. The availability of different types of traffic, accident/incident, injury, and fatality data was also provided by the ONRSR. An ONRSR-reportable occurrence is defined as “An incident associated with railway operations that has, or could have, caused significant property damage, serious injury or death.” (Office of the National Rail, 1951). There is no reporting threshold for an accident or incident to the ONRSR, so the collection of accident data is potentially more comprehensive, but this makes it less comparable to the data that have a specified reporting threshold. In this study, the ONRSR-reportable derailments, collisions, highway-rail grade crossing incidents, obstructions, train fires, and other miscellaneous accidents were considered. Similar to the European nations, data on the type of train these accidents occurred on is limited. Near miss and precursor data were also collected in this study, again not for direct comparison, but for consideration of rail safety data completeness and resolution.

Table 2
Summary of public data sources for the six selected countries.

<table>
<thead>
<tr>
<th>Country</th>
<th>United States (US)</th>
<th>United Kingdom (UK)</th>
<th>Germany (DE)</th>
<th>France (FR)</th>
<th>Sweden (SE)</th>
<th>Australia (AU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publicly Available Data Sources</td>
<td>FRA, BTS, STB</td>
<td>ORR, EUAR, Eurostat</td>
<td>EUAR, Eurostat</td>
<td>EUAR, Eurostat</td>
<td>EUAR, Eurostat</td>
<td>ONRSR, World Bank</td>
</tr>
<tr>
<td>Traffic Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Train-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Freight Train-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Passenger Train-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Non-Mainline Train-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Freight Ton-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Freight Car-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Accident/Incident Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting Threshold</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Accident Type Breakdown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derailment</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Collision</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Highway-Rail Grade Crossing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Train Fire</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Obstructions (Striking Objects in Non-grade-crossing Areas)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hazardous Material Releases</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distinction Between Passenger and Freight Train Accidents</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distinction Between Mainline and Non-Mainline Train Accidents</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Trespasser Incidents</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Railroad Suicide Data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Precursor/Near Miss Data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Consequence Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Casualty Data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Passenger Casualty Data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Workforce Casualty Data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Non-Railroad Casualty Data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Number of Rail Cars Derailed</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hazardous Materials Release Consequences</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distinction Between Passenger and Freight Train Casualties</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distinction Between Mainline and Non-Mainline Casualties</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distinction Between Different Levels of Injuries</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Available Safety Performance Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Train Accident Rate by Train-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Total Train Accident Rate by Freight Ton-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Total Train Accident Rate by Freight Car-Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Total Train Accident Rate by Passenger Miles</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

a Data availability check marks in bold and with underline indicates higher resolution compared to other countries.
b Official Australian rail safety-related data for the most recent five years are incomplete.
c The ORR data do not have a specified reporting threshold, while the EUAR data do.
all accident/incident, and consequence data for each country is indicated by check marks in table. Of all the countries had passenger and freight-train-mile data, freight ton-mile data, and passenger-mile data which were used to normalize train accident rates. The US was the only country that had non-mainline (yard and industry lines) train-mile and freight car-mile data. Overall, the US has the best resolution in their rail safety data, with 28 of 28 indicators available, only missing the distinction between minor and major injuries and precursor and near miss data. Setting the US as the benchmark, the UK data is ranked as “-3” because it has 23 indicators available in their databases, which is three fewer than the US (the benchmark country, best in class). Germany, France, Sweden are all ranked as “-7” as they all have 19 indicators available. Australia is ranked as “-8” because only 18 of 28 indicators are available in their rail safety data, which is 8 fewer than the benchmarking country. The number of data resolution indicators is one of the approaches to evaluate the quality of rail safety data, but the types of those indicators available are also important because they affect whether certain rail safety performance statistics can be derived, and these were discussed next.

All six countries published data on common railroad accident types such as train derailments, collisions, highway-rail grade crossing accidents, and train fires. The US had higher resolution in traffic and accident data because of the ability to distinguish between different train and track types. The US and UK had detailed information regarding obstruction accidents where a train strikes a non-train object such as fallen trees or rocks on the track. The UK had pre-defined categories for obstruction incidents while the US data relied on incident narratives to identify the objects. Accidents involving hazardous materials (hazmat) were available in all countries except Australia. The US data had higher resolution in terms of rail car type, hazmat information, and release consequences. The US is also the only country among those studied that published train accident counts by train type (freight, passenger, MoW, etc.) and track type (mainline, siding, yard, and industry). All countries except the US published near miss and precursor data. The European nations had higher resolution of these data collected and organized by the EURA. Among them, UK has a more complete set of near miss and precursor data due to the ORR’s inclusion of non-potentially higher-risk train accidents (Non-PHRTAs) which are those that have a lower potential for serious consequences (Office of Rail and Road (ORR), 1230). All countries recorded railroad casualties (fatalities and injuries). The US and UK had higher resolution casualty data because the US broke them down by train and track type, and the UK distinguished between different injury levels. Overall, the UK had the highest resolution data on rail injuries and fatalities.

Based on the review of data systems, rail safety performance measures can be developed for each country. Several key measures were selected for this study based on their common usage in evaluating railroad system safety: overall train accident rates, highway-rail grade crossing incident rates, and railroad trespasser and suicide incident rates. Note that the resolution of data will affect the availability of safety performance measures. Because the US had the highest data resolution, therefore it can derive most of the safety performance measures except precursor/near miss rates. Other countries can derive fewer numbers of safety performance measure depending on what types of data are lacking. Safety performance measures that were produced by all countries were compared and discussed in the next section.

2.2.6. Rail safety performance

Train accident rates were derived by dividing the total number of accidents by several metrics of traffic exposure, specifically train miles (Fig. 3a), passenger miles (Fig. 3b), and freight ton-miles (Fig. 3c). Of the six countries, the US had the highest total train miles, most of which are freight train traffic, the largest amount of freight ton-miles, and the least passenger miles. Germany had similar total train-mile traffic compared to the US, but the majority was passenger train miles. Germany and France had the most passenger mile traffic, followed by the UK, Sweden, and the US. In terms of freight ton-miles, after the US, Australia had substantial freight traffic compared to any of the European countries. Freight traffic in Australia and the US is characterized by heavy tonnage transported over long distances and reflects their much larger size compared to European countries. European countries have denser rail networks and train traffic and higher passenger miles compared to the US and Australia despite their smaller land areas. Some other measures of traffic exposure were considered for specific types of accident or incident. For example, number of grade crossings and highway traffic at grade crossings can be used as alternative measures of exposure for derivation of grade crossing incident rates; however, these data were not publicly available for all countries, and therefore could not be used to derive grade crossing incident rates in this study.

Fig. 4a presents the overall train accident rate by total train miles and Fig. 4b presents the accident rate excluding reportable grade crossing accidents. Due to its generality, train-mile-based accident rate is among the most frequently used rail safety performance measures (Wang et al., 2020; Barkan et al., 2003; Liu, 2015; Zhang et al., 2019; Lin et al., 2020). The US had the highest overall rate of the six countries, ranging from 2.54 to 2.83 accidents per million train miles (APMTM), followed by Australia ranging from 1.12 to 1.34 APMTM (for 2016 to 2019 only). Train-mile-based accident rates for Germany, France, and Sweden ranged between 0.19 and 0.34 APMTM. UK had the lowest accident rate, ranging from 0.02 to 0.09 APMTM. When reportable grade crossing accidents are excluded, the accident rates of all countries decrease, but the overall trend and the relative performance among these countries do not change. If the non-reportable grade crossing accidents are included in the US statistic, the US accident rates would be higher than what was shown in the figures. Accident rates based on passenger-miles and freight ton-miles were suitable for evaluating passenger train and freight train accident respectively. These provide different estimates for risk evaluation that may be more useful in some contexts than estimates based on train-miles (Evans, 1997; Liu et al., 2017). Nevertheless, due to the lack of train accident breakdown by train type for some countries, benchmarking accident rates by these two traffic measures was not possible, because if the total number of accidents was used as the numerator, the resulting rates would have been meaningless as a result of the different train traffic compositions discussed above.

Grade crossings are a major safety concern in all countries (Chadwick et al., 2014; Keramati et al., 2020) and reportable grade crossing incident rates per million train-miles (Fig. 5a) and per 1000 grade crossings (Fig. 5b) were calculated. The US and Australia had the highest rates by train-mile measure and both countries’ rates increased during the period studied in this paper.
Fig. 3. Railroad traffic by (a) train miles, (b) passenger miles, and (c) freight ton-miles.

* Australian train mile traffic in 2015, passenger mile traffic in 2019, and ton-mile traffic in 2017 and 2018 are not available.
Train-mile-based grade crossing incident rates for Germany, France, and Sweden were similar, ranging between 0.08 and 0.18 APMTM, followed by UK, which had the lowest rate. Germany had the highest grade crossing rates by number-of-crossing measure while UK still had the lowest rate, with other countries' rates in between. The differences in the ranking of grade crossing rates by the two types of traffic exposures and their implications were discussed in the Discussion section.

Rail trespassing and suicides have also been a problem for many years and over the past decade have emerged as major concern (Wali et al., 2021; Barker et al., 2016), so suicide and trespasser fatality rates per train mile were calculated (Fig. 6a). The US, Germany, and France had higher rates than Sweden, the UK, and Australia. The US suicide and trespasser fatality rate increased over the period studied, Germany’s and France’s decreased, and the UK and Australia evinced no particular trend. While both rail trespassing and suicide refer to the intrusion of people onto railway tracks, they stem from different motivations and therefore may have different risk
characteristics and demand different risk mitigation strategies. Hence, a further breakdown of rail trespassing and suicide fatalities rates were derived (Fig. 6b and c). The US has the highest rail trespass fatality rate while it has the lowest suicide fatality rates among the six countries based on train-mile measure. Germany has a decreasing trend in suicide fatality rate. It is possible that Germany implemented effective strategies to reduce its suicide rate (Sherry, 2016). Previous studies have shown that if the person attempting suicide must overcome obstacles in order to end their life, they are less likely to do so (Sherry, 2016).

Fig. 5. Comparison of rail safety performance on reportable grade crossing incident rates by (a) million train-miles and (b) 1000 grade crossing.
3. Discussion

3.1. Data completeness and resolution

Although all of the countries included in this study had high data availability and transparency, there were differences in data resolution. For example, the finer granularity of casualty categorization in some countries allowed more accurate severity assessment such as Fatalities and Weighted Injuries/Index (FWI), which has been used in various railroad risk analyses (Rungskunroch et al., 2021; Lin et al., 2020; Covey et al., 2010; Sadler et al., 2016; Leitner, 2017).

The lack of accident data broken down by train and track type is an important limitation of some countries’ data systems. Passenger trains and freight trains have different operating characteristics, so their accident characteristics and safety performance will also differ. For example, the overall train accident rate in the US in 2017 was 2.54 APMTM, but if passenger train and freight train accidents rates are derived separately, the former was 0.90 APMTM (Lin et al., 2020), and the latter 2.77 APMTM (Lin et al., 2018). If the overall rate is the only available measure, one cannot quantify these differences thereby obscuring insight into what may be causing the differing trends. Furthermore, knowing the number of passenger train and freight train accidents facilitated the calculation of accident rates based on passenger miles and freight ton-miles, which provided insights into train-type-specific accident characteristics. Meanwhile, the lack of certain traffic exposure measures such as car miles and number of grade crossings limits the resolution of the evaluation and comparison of rail safety performance. In this study, safety performance is mainly measured by train-miles due to its generality and availability for all six countries under review. Depending on circumstances, safety performance evaluated using other measures of exposure can provide more accurate assessment to the safety of a railroad system. This limitation underscores the recommendation of collection and publication of more detailed rail traffic data.

The distinction between the safety performance of mainline and non-mainline operations is also important because their operating environments vary widely thereby affecting accident characteristics (Anderson and Barkan, 2004). The consequences of train accidents on mainlines and in yards also differ. For instance, mainline train accidents often cause onboard passengers and train crew casualties, whereas non-mainline accidents are more likely to involve railroad workers. Unique traffic exposure data for railyards such as the number of railcars processed, and the number of railcar arrivals and departures can be used to evaluate non-mainline accident rates in addition to train-mile and ton-mile traffic. More specifically, accident rates and severity based on number of railcars processed in the yard and number of yard arrival/departure and yard accident counts can be derived, and the resulting yard accident risk can be integrated with mainline risk to conduct network-based rail transportation risk analyses that account for both mainline operation and yard switching activities (Zhao and Dick, 2022). This is especially useful when addressing freight and hazardous materials transportation operational safety. The granularity of accident and traffic data broken down by track type enables more accurate calculation of mainline and non-mainline accident rates. It allows more detailed risk assessments that account for the distinctive operating circumstances between the two, enabling better focused risk mitigation strategies for the two types of operation.

The US is the only country that provides publicly available accident frequency broken down by train and track type. Although other
countries’ non-mainline operations may be less than the US due to less freight-train traffic, it is still beneficial to record these data to identify potential system safety weaknesses. These data will also be useful if there is a change in operating circumstances so that risk estimates for a particular train or track type can be revised to account for the change.

3.2. Accident precursors and near misses

Collection and analysis of precursor data is beneficial for improving safety because it is a proactive approach to identify potential risks that can lead to a train accident (Sangiorgio et al., 2020). European countries that collect and use precursor data effectively reduced train accident rates to a lower level than the US that does not systematically collect and use such data (Fowler et al., 2013). The US has a Confidential Close Call Reporting program where railroad companies can voluntarily join and report near misses, but the data are not publicly available and are collected only for a limited number of participating railroads (Raney et al., 2019). Precursor analysis is even more crucial in railroad systems with low accident occurrence rates because in these systems, train accident data will be sparse, and therefore quantitative and statistically based risk analyses are less reliable due to lack of a robust dataset (Wright and van der Schaaf, 2004). In this case, analyzing precursor and near miss data can identify system weaknesses and underlying causes that could lead to a train accident that would not have been discovered by train accident data analysis.

3.3. Highway-rail grade crossing risk

Grade crossing (level crossing) risk reduction is an ongoing objective for all six railroad systems. The respective roles of trains, roadway traffic, and pedestrians makes the assessment and reduction of such risk more complicated and requires collaboration between railroad and highway agencies. Thus, the availability of comprehensive incident data is indispensable. It is observed that the relative ranking of incident rates for the six countries differed when different traffic measures, i.e., train-mile and number of grade crossings (Fig. 5a and b), were used, and this reflected the characteristics of individual country’s rail network, grade crossings, and rail/highway traffic. The European nations’ grade crossing incident rates were lower than the US and Australia over the years considered in this study on train-mile basis (Fig. 5a). It is possible that some unique system features in the US and Australia lead to higher train-mile-based incident rates (for example, the longer average train length in these two countries due to the greater proportion of freight train traffic leads to impatience of roadway vehicle drivers). On the other hand, some European countries had higher grade crossing rates than the US and Australia by number-of-crossings measure (Fig. 5b). A potential explanation could be that more grade crossings in the European countries have denser train and roadway traffic, thus increasing the per-crossing accident rates. It is noteworthy that the UK has the lowest grade crossing accident rates by both traffic measure, indicating that its grade crossing safety performance is better than other countries. Accordingly, since there are some common system features between UK and other countries, some systematic grade crossing safety improvement initiatives developed by the former can be referenced (Network Rail, 2019; Office of Rail and Road (ORR), 2021). In addition, although all the nations have been working on further reducing the risk, there is no obvious trend of improvement in the years reviewed. Given that most countries collect high-resolution grade crossing incident data, there is potential for a meta-analysis to combine and analyze these data together and pinpoint key factors affecting the risk that may not be captured by analyzing individual countries’ incident data. It is also important to note that additional grade-crossing-related traffic data can be collected to further derive more accurate safety statistics. For example, collecting the number of grade crossing in a rail network allows the derivation of incident rate based on grade crossing density. Comparison of train-mile-based incident rate and crossing-density-based incident rate offers more information for analyzing and benchmarking grade crossing risk among different railway systems.

3.4. Trespasser and suicide incidents

Rail trespasser and suicide risk account for the largest percentage of rail-related casualties in many railroad systems. In recent years more research has addressed this risk to develop mitigation strategies (Wali et al., 2021; Ryan et al., 2018; Zhang et al., 2018b; Beiler et al., 2019). Identifying locations along a railroad track that have a high risk of trespasser and suicide events requires comprehensive collection and analysis of data. Similar to grade crossings, a meta-analysis that leverages multiple statistical results from different countries may provide insights into key factors affecting the risk.

3.5. Impact of high-profile train accidents

While one of the main purposes of collecting train accident and incident data is to obtain sufficient sample size for effective statistical inferences and risk analyses, the lessons learned from few but very severe accidents are also crucial for rail safety improvement. In fact, the occurrence of a severe train accident indicates that there must be multiple flaws in the safety of the rail system, some of them are critical and must be taken care of immediately. Some severe train accidents have led to the legislation of new regulations or implementation of new practices for safety rail operations. From the perspective of rail safety data collection and analysis, very severe train accidents also motivated the recording of additional information on routine basis in the accident/incident database with the aim to identify and remove risk factors that could lead to severe accidents in advance.

Table 3 summarizes selected high-consequence train accidents in the six countries of interest since 2015. The most common types of high-profile train accidents are derailment and collision, with a few grade crossing incidents and hazmat releases. It is worth to note that sometimes certain severity measures may not reflect the true severity level of the accident. For example, in the UK’s train
Table 3
Selected recent high-consequence train accidents for the six countries.

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Location</th>
<th>Type</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>2015</td>
<td>Philadelphia, Pennsylvania</td>
<td>Derailment</td>
<td>8 fatalities 185 injuries About 31 million US dollars in damage</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>DuPont, Washington</td>
<td>Derailment</td>
<td>3 fatalities 57 injuries About 26 million US dollars in damage</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>Hyndman, Pennsylvania</td>
<td>Derailment and Hazardous Materials Release</td>
<td>13 railcars derailed 3 hazmat cars released About 1.4 million US dollars in damage</td>
</tr>
<tr>
<td></td>
<td>2018</td>
<td>Cayce, South Carolina</td>
<td>Collision</td>
<td>2 fatalities 91 injuries About 25 million US dollars in damage</td>
</tr>
<tr>
<td></td>
<td>2022</td>
<td>Mendon, Missouri</td>
<td>Grade Crossing Collision</td>
<td>3 fatalities Multiple injuries About 4 million US dollars in damage</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2016</td>
<td>Plymouth</td>
<td>Collision</td>
<td>48 injuries Damage to both colliding trains</td>
</tr>
<tr>
<td></td>
<td>2020</td>
<td>Carmont, Aberdeenshire</td>
<td>Derailment</td>
<td>3 fatalities 6 injuries</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>Salisbury</td>
<td>Collision</td>
<td>14 injuries Damage to both colliding trains</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tunnel Junction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>2016</td>
<td>Bad Aibling</td>
<td>Collision</td>
<td>12 fatalities 85 injuries</td>
</tr>
<tr>
<td></td>
<td>2022</td>
<td>Garmisch-Partenkirchen</td>
<td>Derailment</td>
<td>5 fatalities 68 injuries</td>
</tr>
<tr>
<td>France</td>
<td>2015</td>
<td>Eckwersheim, Alsace</td>
<td>Derailment</td>
<td>11 fatalities 42 injuries</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>Millas, Arrondissement</td>
<td>Grade Crossing Collision</td>
<td>6 fatalities 23 injuries</td>
</tr>
<tr>
<td>Australia</td>
<td>2018</td>
<td>Port Hedland, Western Australia</td>
<td>Derailment (Runaway Train)</td>
<td>2 locomotives, 245 railcars, and 2 km of tracks were destroyed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

derailment at Carmont, Aberdeenshire, in 2020, there were three fatalities and six injuries. It did not seem to be particularly severe simply by the number of casualties, but this was because at the time of the accident, there were only nine people on the train. The train actually suffered severe impact, and all personnel onboard were either fatally injured or severely injured. To mitigate such effect that may mask the true accident severity by considering a single consequence measure, one may consider multiple types of consequence measures in train accident consequence evaluation, with proper conversion of different types of consequences to make them comparable. Not all countries record details of different types of consequence consistently. For instance, the US records monetary loss for almost all of its train accidents, especially the severe ones, while the European countries do not document such details in all of their accidents. Sweden does not have any high-profile train accidents in the recent years, though its accident rates is not the lowest.

3.6. Railroad operational environment and safety data/performance

The development of railway systems varies country by country, and the resulting varying structures of railroad operations and infrastructure managements under different history and background lead to differences in the construction and maintenance of rail safety data and focuses on rail operational safety and risk management. Table 4 summarizes the operational structure and management themes of railway systems and the primary and recent focuses on rail safety in the six countries. It is observed that most countries’ rail infrastructure is primarily managed by state or federal government, some with small portion managed by private companies. In the US, the infrastructure is mostly owned and managed by private railroads, and they own and operate their own trains while allowing interchanges among different railroad companies in the general railroad systems. In other countries, trains are operated by the state or federal government or private operators who lease trackage accesses from the infrastructure managers. Different managerial structures on railway infrastructure, rolling stock, and train operations may affect the safety performance of the railway system, depending on how the involving organizations are integrated, and the risk management strategies and resources available from individual participants in the rail operation. For example, if the infrastructure of the whole rail network is managed by one entity such as the government, measures implemented to control and mitigate risks can be consistently and uniformly in place throughout the network, but if the rail infrastructure is managed by various parties on different parts of the rail network, they may have different resources, strategies, and focuses so that the actions taken on different track segments to control risk will be different, affecting their safety performance. These may also affect the way rail safety data are constructed, maintained, and integrated. For instance, Australia has recently initiated a National Rail Safety Data Strategy program with the primary goal to integrate various sources of rail safety data into a relevant, consistent and quality data platform (Office of the National Rail Safety Regulator, 2018). When comparing rail safety data and performance, it is important to understand the fundamental differences in characteristics of different railway systems that may affect the interpretations of the results of benchmarking.

From Table 4, it is also observed that different countries focused on addressing different sets of safety issues, with some overlapping topics and some unique and country-specific topics. For instance, railway worker safety, level crossing safety, and suicide and trespasser prevention have all appeared as important safety focuses in multiple countries, representing opportunities for these countries to exchange information and identify best practices that can be adapted to respective countries (and other countries in the future if they would like to focus on these issues). There are some safety issues that gain more attention in certain countries due to their operational environments. For example, because of the significant amount of traffic and importance to the economics, safety of rail hazardous materials transportation has been an important risk analysis topic in the United States, and specific research were conducted, and
<table>
<thead>
<tr>
<th>Country</th>
<th>Operational and Management Structure</th>
<th>Primary and Recent Safety Focus</th>
</tr>
</thead>
</table>
| United States | - Most railroad infrastructure owned and maintained by private companies  
- Freight trains operated by private companies; passenger trains mostly operated by government or government-subsidized agencies.  
- Freight trains are the dominant traffic type compared to passenger trains                                                                                                                                                                                                                                                                                                                                                                           | - Rail hazardous materials transportation safety  
- Railway worker safety  
- Positive Train Control (PTC)  
- Trespassing detection  
- Operation Life Saver (level crossing)  
- Rail suicide prevention                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| United Kingdom | - The infrastructure network is owned and managed by a public company (Network Rail). Sections of the network are franchised to private operators in contracts of around 7–10 years. Franchises pay premiums to operate main lines where demand is high; franchises receive subsidies to operate low-demand regional lines. Passenger traffic is dominant, and passenger trains have priority over freight trains in the timetable.                                                                                                                                                                                                                                                   | - Improvement to inspection, maintenance and renewal of drainage assets  
- Targeted vegetation management  
- Measures to reduce the risk of underbridge scour at higher risk sites  
- Review of solid state interlocking signalling assets Installing overlay systems at passive crossings, replacing telephones and whistleboards where possible  
- Renewing automatic level crossings with safer designs  
- Development and rollout of automatic full barrier crossings  
- Red light safety cameras  
- Targeted level crossing safety awareness campaigns                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Germany      | - Infrastructure is owned and managed by a public company (Deutsche Bahn, DB). Operations are provided both by DB own companies, or private companies. Traffic is 50% passenger, 50% freight.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | - Awareness-raising campaign for the prevention of rail accidents and for more civil courage, to raise passengers’ awareness of the need to use care in train stations and in particular on platforms  
- Implementation of the National Noise Protection Act  
- Advancing the digitalization of the network, for example through the comprehensive rollout of European Train Control System (ETCS), digital interlockings and digital rail operations                                                                                                                                                                                                                                                                                                                                                                  |
| France       | - Infrastructure is owned and managed by a country-owned company (SNCF). Infrastructure is managed by SNCF Reseau. Operations are done mainly by SNCF. Regional trains (TER) are managed by local authorities under contracts with SNCF. Almost all high-speed rail services are run by SNCF. The HSR between UK and France is operated privately by Eurostar. Recently, private operators began to operate services in the country (Thello).                                                                                                                                                                                                                   | - Prevent risk on the workplace for railway workers  
- European Railway Safety Culture Declaration                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Sweden       | - Infrastructure is owned and managed by a country-owned company (Trafikverket), which contracts private companies for maintenance. There is a state-owned operating company (SJ) which competes with private operators through tendering of franchises and open access.                                                                                                                                                                                                                                                                                                                                                              | - Network-wide introduction of European Rail Traffic Management System (ERTMS)  
- Remote control of railways and the introduction of the Future Railway Mobile Communications System (FRMCS)  
- National Rail Safety Data Strategy  
- Fatigue, drug and alcohol management  
- Level crossing safety  
- Track worker safety  
- Suicide and trespasser prevention                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Australia    | - Infrastructure is managed mainly at state level with connections maintained by the federal government with a few privately owned track sections. The infrastructure managers do not operate revenue trains. Instead, they provide track access to train operators. Passenger and freight train operations are provided by multiple public and private companies who purchase track access from the infrastructure managers.                                                                                                                                 |

C.-Y. Lin et al.
preventive measures were identified and implemented to address such risk (Lin et al., 2022; Iranitalab and Khattak, 2020; Kang et al., 2022).

3.7. Cross-referencing opportunities for improvement

The essence of benchmarking is the process of identifying the highest standard of excellence, and then making the changes necessary to achieve those standards – commonly known as ‘best practice’ (Bhutta and Huq, 1999). Through the standardized benchmarking framework and procedure presented in Fig. 2, rail safety-related data systems and high-level safety performance of the six countries were analyzed, and their strengths and weaknesses identified. Following is a summary of findings from the six countries reviewed.

- The US had the highest resolution of publicly available railroad safety-related data, attributing to the various FRA databases. These data allowed breakdown of traffic and accidents by train type, track type, and other accident features such as weather, employee hours, method of operation, time of day, etc. This high-resolution data has been leveraged to conduct various quantitative rail safety analyses and train accident prediction modeling (Wang et al., 2020; Barkan et al., 2003; Liu, 2015; Zhang et al., 2019; Lin et al., 2020; Liu et al., 2017).
- The US had the most comprehensive information on hazmat in their traffic and accident data. FRA accident data records if a train carried hazmat, the number of hazmat cars derailed, damaged, and released, and name of the products involved. These data have enabled quantitative risk analyses of rail hazmat transportation (Barkan et al., 2003; Iranitalab and Khattak, 2020; Kang et al., 2022).
- Australia, the UK, and European countries systematically collect and analyze precursor and near miss data. This enables prior detection of hazardous activities or situations that may lead to train accidents (Muttram, 2002; Quigley et al., 2007). Such information is beneficial because potential problems or trends may be foreseen before a more serious accident occurs. US railroad safety improvement efforts would benefit from development, deployment, and use of a precursor and near miss data system; any institutional barriers to such a system should be overcome.

Benchmarking permits each country’s rail system and regulator to evaluate other countries’ rail safety-related data systems and use these to inform potential improvements to their own. For example, the analysis described in this paper suggests that various traffic exposure metrics, more granular breakdown of train accident statistics, and near miss data should be consistently and systematically collected. These would make more accurate and informative rail operational risk assessment possible. Moreover, risk mitigation measures that have been proven effective in one country for a specific safety problem may, with suitable modifications, be adapted to another’s rail system. For instance, experiences in reducing broken rail and other rail and switch defects due to heavy haul freight operations in the US can be referenced by other countries if heavy haul freight operations are to be introduced. In another instance, documentation and management of near miss and precursor data systems developed by European countries can be referenced by the US to develop their own risk management strategies. As for some safety issues that are already focused on by multiple countries such as rail suicide, risk preventive measures from different countries (for example, the “There Is Always hope” campaign in UK, the “Global Railway Alliance for Suicide Prevention” organized by the US) can be further benchmarked to identify best practices so that each country can benefit from successful experiences from other countries and enhance their risk mitigation measures.

It is not the authors’ intention to criticize or call out system weaknesses in different countries’ rail safety performance. Rather, this research is intended to identify and understand relative strengths and weaknesses in safety data systems and the implications for what can and cannot be measured using them. It is hoped that this will provide insight into the most effective safety improvement strategies for all countries’ rail systems in general, with particular focus on best practices to enhance rail safety data systems enabling more effective safety improvement strategies.

All of the findings described here are based on publicly available data. It is possible that some non-publicly available or higher-resolution data exist for internal use by railroad operators or government agencies. Although these data may provide internal utility, the consequence of less accessible data is reduced benchmarking efficacy. Furthermore, there is potentially additional statistical and inferential power that could be achieved if multiple data systems could be leveraged such as a meta-analysis.

Benchmarking is a process to establish the basis for creative breakthroughs, rather than competitive evaluation and numerical analysis, espionage, or data theft (Bhutta and Huq, 1999). Most countries publish little or no data (Blumenfeld et al., 2023) thereby restricting the ability to benchmark at all and curtailing opportunities for safety improvement that might otherwise be possible. Sharing safety-related data should be seen as beneficial for all parties as cooperative, knowledge-based approaches to safety improvement are more efficient than individual efforts in which each country must learn on its own, often through difficult and costly experience.

4. Concluding remarks

This paper presents a framework for benchmarking railroad safety data and safety performance among different railroads’ systems. This framework consists of a standardized benchmarking wheel structure specialized with railway data system elements. Six countries with high data availability and transparency were selected to conduct a comparative analysis. Twenty-eight resolution indicators were assessed for all six countries, and the one with the highest resolution was set as the benchmark. The strengths of each country’s publicly available railroad safety-related data sources were then identified. Prior to this research, there has been no studies attempting to
analyze the resolution and quality of rail safety databases. Through the benchmarking framework presented, the results showed the resolution of different types of rail safety-related data differed among the six countries, and the relative advantages and disadvantages of different countries’ data system were discussed. Since data resolution is an important element in evaluating rail safety data and affects the availability and accuracy of rail safety statistics and subsequent evaluation of rail safety performance, a more in-depth and dedicated research focusing on assessing rail safety data resolution is called for.

Meanwhile, through the benchmarking among the six countries, this paper presented the best practices of rail safety data collection by combining the advantages of individual country’s rail safety databases. The following data are considered crucial to derive key safety performance statistics and should be systematically collected, maintained, and publicized.

- **Rail traffic data**: train-miles, car-miles, (freight) ton-miles, passenger miles, number of grade crossings, hazardous materials traffic, with data capable of being broken down by train type and track type, rail yard switching traffic (rail car processed and departure/arrival).
- **Rail accident data**: derailments, collisions, grade crossing incidents, obstructions, train fire, trespassing, suicide, hazardous materials releases, precursor and near miss, with data capable of being broken down by train type and track type.
- **Rail accident consequence**: number of rail vehicle derailed, number of hazardous materials cars released, fatalities, different levels of injuries, monetary loss, service disruption (train delay), with data capable of being broken down by train type and track type.

High-level rail safety performance of these countries was analyzed and compared using the data. The UK had the lowest overall accident and grade crossing incident rates among countries studied. They have several best practices that help them achieve this level of safety that can be referenced and adapted by other countries. European nations have used the collection and analysis of precursor and near miss data to reduce train accident risk. This study is the first to benchmark different countries’ rail safety data systems using publicly available sources internationally. By identifying and adapting the advantages of each country’s railroad safety data system and risk mitigation measures to other countries, the overall safety performance of individual railroad system can potentially be improved more effectively and efficiently. The authors believe that sharing rail safety-related data such as accident reporting criteria, frequency and severity of train accidents, precursor and near misses, and rail traffic are beneficial. Another benefit to provision of publicly available data is that it permits more robust datasets for risk assessment. Techniques such as meta-analysis can be leveraged to improve the accuracy and strength of statistical analysis and risk evaluation. Meta-analysis is a statistical technique that combines research results for the same research subject to obtain a more accurate and less biased outcome (Hedges, 1992). It is a systematic review procedure that has been proven effective in various research fields such as medical science, agricultural research, social science, etc. (DiMatteo et al., 2002; Challinor et al., 2014; Huang et al., 2009; Melo et al., 2013), and is a promising method for rail safety and risk analysis. More advanced data analytical tools and techniques can also be applied to conduct quantitative risk assessments for the rail systems if complete and high-resolution rail safety data become reliably available. For example, artificial neural network (ANN) can be used to derive railway accident frequency and casualty prediction based on high-resolution train accident data (Lim, 2023). Machine-learning based methods can be used to assess key influencing factors for the financial loss of railway accidents (Dhingra et al., 2022) and accident prediction (Meng et al., 2022). Zero inflated negative binomial and empirical Bayes method can be used to analyze level crossing incident prediction given detail recording of historical level crossing incident data (Mathew and Benekohal, 2021). The benchmarking framework developed in this research can be applied to railroad systems in other countries, and the benchmarking concept can be used for other transportation safety data collection, analysis, and comparison.

**Author contribution statement**

The authors confirm contribution to the paper as follows: study conception and design: Chen-Yu Lin, Christopher P.L. Barkan, Anson Jack; data collection: Chen-Yu Lin, Theodore Gerstein; analysis and interpretation of results: Chen-Yu Lin, Christopher P.L. Barkan, Theodore Gerstein, Anson Jack, Marcelo Blumenfeld, Usman T. Abdurrahman; draft manuscript preparation: Chen-Yu Lin, Theodore Gerstein, Christopher P.L. Barkan, Anson Jack, Marcelo Blumenfeld, Usman T. Abdurrahman. All authors reviewed the results and approved the final version of the manuscript.

**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.

**Acknowledgment**

The authors thank the Birmingham-Illinois Partnership for Discovery, Engagement and Education (BRIDGE) for funding this research. The first and third authors also thank the Association of American Railroads (AAR) for their financial support.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrtpm.2023.100384.