SPECIAL REPORT 323:
IN-SERVICE PERFORMANCE EVALUATION OF GUARDRAIL END TREATMENTS

TRB-SASP-14-05
In-Service Performance of Energy-Absorbing W-Beam Guardrail End Treatments: Phase 1

Critical Review of Methodologies for Evaluating In-Use Safety Performance of Guardrail End Treatments and Other Roadside Treatments

BRIAN WOLSHON
Louisiana State University

ANURAG PANDE
California Polytechnic University, San Luis Obispo

Paper prepared for the
Committee for the Study of In-Service Performance
of W-Beam Guardrail End Treatments, Phase 1
Transportation Research Board

2017

Critical Review of Methodologies for Evaluating In-Use Safety Performance of Guardrail End Treatments and Other Roadside Treatments

BRIAN WOLSHON
Louisiana State University

ANURAG PANDE
California Polytechnic University, San Luis Obispo

EXECUTIVE SUMMARY

In-service Performance Evaluation (ISPE) is a key component of the safety management process for roadside safety equipment installation. This study reviewed road safety literature as well as studies from other fields (with epidemiological studies being most relevant) to identify methods that have the potentially to be applied by state agencies to enhance their ISPE process for guardrail end treatments. The methodologies reviewed here are most relevant to the analysis of data within broader ISPE process. However, a complete ISPE process would involve a much more comprehensive set of activities, including continuous gathering and maintenance of crash details, location information, and equipment inventory data elements.

This report is organized into three sections that encompass the

- Motivation and objectives of the study;
- Review of literature on key elements of
  - existing ISPE process and its goals,
  - different research designs and analysis techniques for observational data analysis, and
  - practices of rare-event analysis from other fields of study;
- Key findings and conclusions from the review.

Among the key findings of this study were:

- There has been an extensive and varied history of activity in the US and Canada, led primarily by State Departments of Transportation
- Practices vary from extensive and systematical practices, to quite informal activities
- Geographic differences in practice were also evident
- Analysis objectives and methods are influenced by the cost, time, and effort to collect data to implement the methodology
- Subtle nuances in the safety differences of different vehicle types and medical databases could be taken into account to enhance future ISPE efforts
- Roadway crash databases and other sources of road/crash characteristics can provide valuable input to the ISPE process

- A wide range of statistical data acquisition techniques, databases, and analysis techniques (with varying levels of complexity, cost, effort and resulting utility) have been explored in prior research
- Rare event identification, assessment, modeling, forecasting, and classification techniques in other fields generally showed limited potential for application in the ISPE of guardrail end treatments
- As capabilities to monitor infrastructure and archive safety data increase, it may be possible to apply sensors to guardrail for continuous in-service monitoring.

INTRODUCTION

Every year, about 32,000 people die in motor vehicle related highway crashes in the United States (*1*). Of these, approximately 20 percent are the result of single-vehicle, run-of-the-road collisions with fixed objects like trees and utility poles. The reasons for these crashes are varied, but are often related to poor weather conditions; excessive speed; and driver impairment and distraction; among many others.

Early research by the General Motors (GM) Corporation combined with decades of observational experience showed that the removal of fixed object hazards and the flattening of side slopes within the roadside area adjacent to the shoulder had significant safety benefits (*2*). In fact, GM research suggested that up to 80 percent of run-off-the-road crashes could be prevented by incorporating a 30 foot recovery area free of obstacles and steep slopes. It was from this work that the idea of clear zones was born. Clear zones are roadside areas free of hazards and a with flat, traversable side slopes that permit drivers to regain control of their vehicles and steer back to the road or bring their vehicles to controlled stops without encountering a hazardous condition. Today, clear zones are a fundamental component of safety conscious road design and recommendations for their design are found in the American Association of State Highway and Transportation Officials (AASHTO) *Roadside Design Guide* (*2*).

Obviously, however, it is not always practical or even physically possible to incorporate a clear zone into all designs. Any number of practical considerations can preclude the removal, relocation, or redesign of fixed object hazards and flattening of side slopes in the roadside. When this occurs, one design alternative may be to incorporate a longitudinal barrier beyond the roadway shoulder. These barriers act by containing and redirecting errant vehicles away from the hazard.

Unfortunately, roadside barriers also present a significant hazard to vehicles and drivers in and of themselves. However, they are regarded to be a preferred alternative to prevent vehicles from impacting a fixed object or traversing a slope with a high potential for vehicle rollover or an inability to safely steer back to the road. While designed to provide lateral deflection and/or vertical redirection in side angle impacts, roadside barriers can present a significant safety hazard when impacted end-on. To lessen this threat, a variety of crashworthy end terminal designs have evolved over decades to limit the danger posed by end-on impacts. This process has resulted in modern designs that seek to provide tensile stiffness to offer redirectional longitudinal rigidity for lateral side impacts, while also reducing abrupt decelerations for end-on impacts.

Despite the demonstrated benefits of end treatments, their designs continue to advance. A key part of the advancement process, and the one that is the focus of this project, is gaining an understanding how they perform under field conditions and how their performance may change over time. Although the performance characteristics of all roadside barriers and end terminals are rigorously and systematically tested in field test experiments, such research occurs under tightly controlled conditions that specify specific construction, installation, and crash (speed, impact point, angle of approach, vehicle weight, etc.) conditions. Once installed in the field, however, these details can vary widely from these testing specifications and the specific characteristics of any particular crash can vary nearly infinitely. Complicating matters even further, it is recognized that the performance of barriers can also vary over time as materials age and road conditions alter with maintenance and deterioration.

This paper documents the findings of a review of literature related to the methods used for and knowledge gained from in-service performance evaluations (ISPE) of guardrail end treatments. In broad terms, in-service performance evaluations are processes that are used to assess the behavior and characteristics of engineered systems in in-use settings to gauge how they are performing or may be expected to perform in the future. These evaluations are often carried out on informal ad-hoc bases, frequently in response to a sequence of crash events involving a particular device, even though processes for continuous monitoring on an ongoing basis are available in the literature [e.g., (*3, 4*)].

In this research, several aspects of ISPEs are described, most notably approaches (or lack thereof) used by transportation agencies throughout the US to conduct ISPEs of end treatments and, more generally, methodologies used to conduct in-service safety evaluations of all roadside safety features. The methodological review ranges from more commonly-used injury severity analysis techniques to methods used for rare-event prediction and analysis in other fields of study. A key aspect of the latter, is an assessment of the extent to which these methodologies may be adapted for application of ISPEs of guardrail end treatments. To help illustrate the utility of range of approaches reviewed in this study, a two-dimensional matrix is included at the end of this report (See Table 3) to summarize key aspects of the methods.

Why ISPE?

Similar to all road safety equipment, barrier end treatment designs are required to go through a series of crash tests prior to their use in the field. While the crash tests are necessary to establish the expected operational performance characteristics of these devices; it is also recognized that actual performance characteristics could vary due to location-specific conditions and environmental effects over time. For example, a barrier may be constructed on shoulders and embankment slopes different from initial crash testing, struck at different angles by vehicles involved in real crashes, and/or subjected to the effects of settlement (*5*). Inconsistent and incorrect reporting of crashes to the police may also not accurately describe the events of a crash involving an end treatment. As a result, practices to monitor barrier end treatment performance in the field, over time, have emerged. In broad terms, these practices are known as “in-service performance evaluations” (ISPE).

The *Manual for Assessing Safety Hardware (MASH)* (*6*) notes the importance of ISPE for roadside devices including guardrails by emphasizing that crash testing “...must be viewed as a necessary, *but not sufficient*, condition to indicate that a feature would perform satisfactorily under real-world conditions. (Emphasis added)".

According to NCHRP Report 490 (*5*) the purpose of in-service evaluations of roadside features is to:

> "…determine how such devices perform under field conditions. Performance, in this context, includes knowing the number, severity, and proportion of people injured in collisions involving the roadside feature; installation and maintenance problems; and the collision, installation, and repair costs associated with the feature."

These evaluations are also expected to inform the formulation of policy and contribute to the development and ongoing growth of performance databases that can be repositories of historical information for use in comparative studies. As further stated in NCHRP Report 490 (*5*):

> "If these performance measures are known, designers and policy makers can optimize the safety benefit obtained by installing the most appropriate roadside features in the most hazardous locations. In-service performance evaluations are the best source of information about injury severity and installation, maintenance, and repair costs. Without good in-service performance information, it is difficult to perform meaningful cost-benefit analyses."

Objectives

The overarching goal of this review, solicited by the Guardrail Performance Evaluation Committee (*7*) of the Transportation Research Board, is to support State transportation agencies in developing procedures to conduct ISPEs of guardrail end treatments. Specifically, the Committee would like to provide these agencies with a starting point so that they are able to, systematically and effectively accomplish one or more of the following tasks:

- Develop data capture protocols for in-service performance evaluation;
- Compare the effectiveness of various guardrail end treatment designs in minimizing injury severity in a crash, while considering factors aside from design that might affect this outcome, such as road conditions, vehicle type, speed, and crash dynamics. Assessment may include quantitative and qualitative elements. The results would aid highway agencies in selecting end treatment designs to be installed at particular locations;
- Evaluate the impact of installation and maintenance practices and environmental conditions on in-service safety performance;
- Estimate benefits and costs of a program of replacing installed end treatments with end treatments that perform better;
- Provide data to support the improvement of end treatment designs; and
- Refine crash test procedures such that it better reflect the in-service performance;

While not intended to provide definitive recommendations/guidance for practice, this report is intended as informational background to support future state agency decision-making and policy-development relative to the ISPE objectives listed here.

Overview of the Report

Primarily, this report provides a comparative review of studies that have been conducted for the purpose of in-service performance evaluation (ISPE) of roadside safety devices (not limited to guardrail end treatments) as well as analogous applications in the other fields. The research question addressed by the evaluation studies is also discussed along with the methodology and results of the evaluations.

The review was conducted to include three broad categories of research, including:

- Studies and state department of transportation reports related with in-service performance evaluations of specific roadside safety equipment including but not limited to guardrail end treatments. This part of the review served to understand the goals state agencies have created for the ISPE process for roadside equipment.
- Evaluation studies that focus specifically on statistical analysis of safety data were reviewed for methodologies that could support the goals of an ISPE process. These studies included crash injury severity analysis as well as before/after evaluation using quasi-experimental design approaches. The review further organizes each study according to a taxonomy of evaluation goals and study designs. The relationship between evaluation goal and appropriate study design is also discussed.
- The final area was a review of data-driven approaches to rare event (e.g., fatal crash) modeling in transportation engineering/safety and other fields of potential interest.

Based on the findings of the review, an attempt was made to summarize and comparatively illustrate methods (and potential approaches) for conducting ISPEs into a single two-dimensional matrix defined by i.) Objective of the evaluation and ii.) Cost/effort involved in collection of the data required by the approach. The idea of the table (See Table 3) was to distill the results of this review such that the TRB committee would be able to convey, as clearly as possible, the most appropriate study design relative to an evaluation objective as well as the data collection costs/effort to the state DOTs. It is thought this knowledge would most-effectively support state agencies' efforts to improve their ISPE practices in the both short- and long-term time horizons.

LITERATURE REVIEW

The review effort began with NCHRP Report 490 (5). This watershed study established ISPE procedures for traffic barriers. As part of this report Ray et al. (5) reviewed the literature to identify past and then current in-service evaluation studies to identify methods that had been effective. Among these were in-service evaluations planned and performed in portions of the States of Connecticut, Iowa, and North Carolina. The pilot studies in the three states demonstrated that ISPEs can yield useful information about the field performance of roadside features that cannot always be revealed in standardized experimental crash tests. Among the conclusions noted in NCHRP Report 490 was that in-service performance evaluations should be integrated more fully into the overall cycle of design, test, and evaluation of roadside hardware.

Essentials of In-service Performance Evaluations

According to Schalkwyk et al. (*3*), who devised and executed a preliminary ISPE process for roadside safety equipment for the State of Texas, the main purpose of the in-service performance evaluation (ISPE) of roadside safety features is to determine:

- "How such devices perform under field conditions, including the vehicle crash experience involving the roadside feature;
- Potential installation and maintenance problems; and
- The collision, installation and repair costs associated with the feature."

Schalkwyk et al. (*3*) also noted the following reasons for state agencies primarily relying on crash tests alone for the safety evaluation of roadside safety devices:

- "no 'formal process' has been established to conduct the evaluation;
- collecting and analyzing the data require a significant commitment of manpower; and
- lack of good, sustainable working relationships among police agencies, area engineers, and maintenance personnel."

This was also emphasized by Ray et al. (*8*) who noted that an essential consideration for conducting analysis to support ISPE is that crash data are reported for various states from various sources using different data collection techniques. It implies that ISPE processes would benefit from statewide standardization of crash data collection methods.

The Arizona Department of Transportation (ADOT) has proposed an in-service evaluation program that consists of four different subsystems that complement each other. This framework attempts to address the causes outlined above. These subsystems are continuous monitoring, supplemental data collection, in-depth investigation, and new product evaluation defined as Levels 1 through 4 respectively (*4*).

- *Level 1*: A continuous monitoring system is part of Level 1 evaluation. In the Level 1 evaluation, a master database is created by linking the following data elements: crash data, maintenance data, highway and traffic data, and the roadside feature inventory.
- *Level 2*: Supplemental data, beyond Level 1, for this level includes:
  - Field data on the roadway, the roadside, and the selected roadside safety feature, and
  - A manual review of the hard copies of the crash reports as completed by the police.

As will be noted in the literature review later, the manual review of crash reports and field data are critical to certain analysis techniques depending on the research question.

- *Level 3*: An in-depth investigation takes place at Level 3 by conducting detailed studies of selected accidents involving a particular set of devices to assess how the particular roadside safety feature performed. This level of detail is necessary if the objective of the evaluation involves determining the mechanical cause of failure. It may involve an on-call team that can investigate (via photographs etc.) the crash before the roadside equipment is repaired/restored.

- *Level 4*: At Level 4, a new product evaluation subsystem is implemented. At this level, potential installation and maintenance-related problems associated with new roadside safety features are targeted.

These levels are sequential in that Level 3 and 4 require significant amount of effort in gathering the data while Level 1 involves linking of the databases most state agencies already gather in some form. Levels 3 and/or 4 subsystem for ISPE can involve roadside systems equipped with sensory devices that can provide information regarding significant impact. Sensor-equipped roadside devices can also support analysis primarily conducted with Level 1 and/or Level 2 information subsystems by potentially ameliorating the problem of under-reporting of non-injury crashes. However, in practice, the under-reporting issue is dealt with, if at all, through application of statistical analysis techniques (addressed later in this paper).

Gabler et al. (*9*) categorized the roadside barrier field experience evaluations into two categories: 1.) in-service evaluations, and 2.) crash studies. An in-service evaluation is a study to evaluate the field performance of a particular roadside safety device. A crash study utilizes crash data (not necessarily involving only the same device) to investigate the performance or relative performance of one or several roadside devices. In this regard, in-service evaluation may be aimed at examining performance in terms of durability and not just in terms of safety. The focus of this work is, however, on safety performance evaluation methodologies and the studies reviewed in the subsequent sections reflect that focus.

Guardrail End Treatment Performance Research

A comprehensive review of ET-PLUS end treatments by the FHWA Task Force found real-world conditions to vary widely from crash testing scenarios (*10*). FHWA conducted a broad search for data on crashes involving extruding w-beam guardrail terminals with the focus on the ET-Plus 4-inch device. FHWA conducted an initial screening of the 1,231 cases received to determine whether they contained sufficient detail to evaluate:

- the type of terminal involved in the crash;
- the role the terminal played in the crash; and
- the performance of the terminal during the crash.

Based upon this initial screening, 161 cases were selected for detailed review and analysis. It was noted, however, that this set of 161 cases assembled by the Task Force was not a representative sample of terminals in-service or a representative sample of terminals that were struck. Despite the lack of representative sampling, the Task Force review revealed performance limitations for ET-PLUS end treatments in two broad areas: 1.) impact conditions and 2.) installation conditions. It was also noted that the performance issues in these two areas may not be limited to the ET-Plus 4-inch devices. The Task Force recommended that comprehensive in-service performance evaluations of guardrail terminals be conducted at the national and State levels.

Agent (*11*) examined the performance of ET2000/ET-PLUS in 135 collisions and were able to review 80 crash reports in Kentucky. Performance was judged to be ‘proper’ in 70 of 80 collisions (88%) in 2004 Kentucky report. The assessment of ‘proper’ vs. ‘improper’ *was not* based on impact severity measured using parameters such as guardrail damage, injury severity, and impact severity. Instead, it was based on whether the end treatment performed as designed. The reason for this choice was that the former may be affected by factors irrelevant to the guardrail performance (e.g., vehicle size and seat-belt use among others). Data collection for the evaluation involved visual inspection of either the damaged guardrail or photographs of the damage and manual reading of crash reports. It was reported that the results of the investigation matched closely with a similar evaluation carried out in 1997.

Esligar and Hildebrand (*12*) examined the performance of the guardrail end treatments in collisions in New Brunswick, Canada from 2007 through 2010. The study involved two energy-absorbing guard rail terminals (EAGRT) systems; the ET-Plus and the SKT-350. While these devices absorbed significant amount of energy it was observed that many of the collision configurations were outside the boundaries defined by NCHRP Report 350 and MASH (Manual for Assessing Safety Hardware). It was also noted that this technology was brought by the manufacturers in the southern United States and these systems may not perform as desired due to significantly different environmental conditions (most notably, snowfall). This conclusion suggested that agencies need to be careful with transferability of ISPE findings from one location to the other. According to the taxonomy of studies noted by Gabler et al. (*9*), the research by Agent (*11*) and that of Esligar and Hildebrand (*12*) would fall into in-service safety evaluation of guardrail end treatments due to their focus on specific device’s performance.

Based on the studies reviewed in this category of analysis, it is suggested that, beyond the basic crash and highway design data, relevant crash reports and inventory of the end-treatment locations are critical elements of the data needed to make any inferences about in-service performance (Levels 1 and 2 noted in the ADOT report reviewed earlier). It is also suggested that a proactive ISPE protocol should also include an On-Call Investigation Team that can assess and evaluate (via photographs etc.) the characteristics and conditions of a crash before the end treatment is repaired/restored (Level 3).

Table 1 summarizes the key information required for ISPE process identified by the ADOT protocol. A key item of note with regard to the ADOT protocol was that it was developed for all roadside features and, as such, does not list the specific data elements for guardrail end treatments. Based on this, it should be expected that desired data elements will vary depending on the agency practices for barrier placement and end treatment design. The documentation of these practices is usually available at State DOT web portals. Two such examples can be reviewed on websites of the Maryland and Pennsylvania State Departments of Transportation (e.g.,(*13, 14*).

TABLE 1 Key Information Needed for Safety Evaluation as Part of the ISPE Process

Information Category	Variable	Measurement Detail/Source
Collision Information	Crash location	GPS Coordinate/Mile posts/Intersecting Road from Crash Record/Report
	End-treatment damage	Crash Record/Report
	Vehicle size	Crash Record/Report
	Impact severity/angle	Crash Record/Report
	Vehicle action after impact	Crash Record/Report
Location Information	Exposure	Length/ADT at the location
	Guardrail placement (e.g., median, roadside, bridge)	Roadside equipment database
	End-treatment configuration (e.g., ET2000, ET-Plus, SRT-350 etc.) and characteristics (e.g.,	Roadside equipment database

It should also be noted that that the studies reviewed in the above-documented discussion include the few studies that address the in-service safety performance of the guardrail end treatments, specifically. This information is insufficient to learn about the methodologies that can be used to make safety related inferences as part of the ISPE process. To enhance this review from a methodological perspective, the following section included a review of studies that aimed at examining performance of other roadside devices.

Evaluating Safety Performance Through Crash Analysis

The methodologies employed for evaluating safety performance using crash data include cross-sectional studies (SPFs), quasi-experimental before-after evaluation based on crash experience (using Empirical Bayes[1] Method (*15*)), and ordered probit/multinomial logit models for injury severity outcomes. Some of the studies examining injury severities also employed case-control designs. This section provides an overview of study designs used for analysis in prior research. Some of the discussion also comes from epidemiological studies. It should be noted once again, that crash analysis studies reviewed in this section will be of interest to reader not just in terms of their findings, but also for insights into the methodologies that were employed to conduct the analyses. The studies that are included here have also been selected to demonstrate research undertaken to address a diverse set of research questions and, as such, methodologies; since the latter depends on the former.

Modeling Approaches and Study Designs

Safety performance of devices may be examined based on injury severity outcome analysis or from counts/rates of crashes before and after installation of the devices. The development of a severity model begins by *conditioning on the event.* This means it is assumed that a crash has

[1] "Empirical Bayes Methodology" is used to estimate the safety performance of a roadway location (e.g., segment or intersection) by combining evidence from crash history of that site with Safety Performance Function (typically negative binomial regression models) generated from similar sites.

occurred and the models do not predict the probability of a crash occurring (*16*). Figure 1 depicts the nested structure typically used for injury severity modeling. Level 1 (i.e., the probability of roadside object crash) is not of interest in injury severity analysis.

Typically, the assessment of likelihood of crash occurrence in the context of in-service performance evaluation is not addressed through the nested structure due to rarity of crash events compared to the potential of having a crash at all. It is also impossible to conduct a randomized controlled experiment, considered the "Gold Standard" in epidemiological research, for in-service performance evaluation. Hence, researchers typically have to rely upon observational studies and/or quasi-experimental before-after evaluation (using Empirical Bayes method) for assessing the impact of traffic safety improvements. Prior to reviewing the crash analysis studies that use these methods, a brief overview of these study designs is provided below.

Case Control Vs. Cross-Sectional Studies

Observational studies are study designs in which no intervention is made (in contrast to an experimental study). Rather, such studies provide estimates and examine associations of events in their natural settings without recourse to experimental intervention (*17*). Cohort studies, cross sectional, and case-control studies are collectively referred to as "observational studies." In this section, cross sectional and case-control study designs, two types commonly used for in-service safety performance evaluations, are summarized.

Broadly, cross-sectional studies[2] are used to determine the prevalence of some occurrence. They are relatively quick and easy to perform, but do not allow the distinction between cause and effect. An example application of cross-sectional study designs was the in the development of safety performance functions (SPFs) used in the Highway Safety Manual (*18*). Case-control

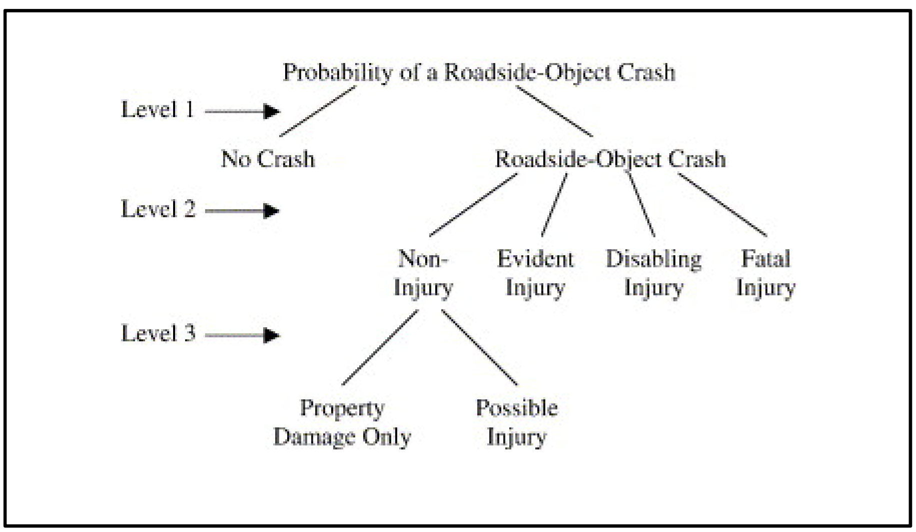

FIGURE 1 Nested structure of crash likelihood and injury severity modeling (*16*).

[2]"Cross-sectional design" involves analysis of data collected from a population, or a representative subset, at one specific point in time (e.g., analysis of total or severe crash frequency on all freeway segments with a particular guardrail end treatment). The units of analysis for these assessments would be the freeway segments.

studies[3] compare separate groups, retrospectively. Effectively, they seek to identify possible predictors of outcome and, as such, are particularly useful for studying rare outcomes (*17*). The case-control may serve as an efficient (in statistical terms) design to identify factors associated with relatively rare severe/fatal crashes compared to non-injury crashes. The efficiency of case-control design as compared to cohort study is detailed by Rothman (*19*).

According to Rothman (*19*) the source population is the population that gives rise to the cases (e.g., severe crashes) included in the study. If a cohort study was undertaken in the context of end-treatment safety evaluation one would define the exposed and unexposed cohorts (all relevant segments with and without a particular end treatment design) and from these populations to obtain denominators for the incidence rates or risks that would be calculated for each cohort. One would then identify the number of cases occurring in each cohort and calculate the risk ratio for each. In a case-control study, the same cases (typically all of the severe crashes) would be identified and classified as to whether they belong to the exposed or unexposed cohort. Instead of obtaining the denominators for the rates or risks (to estimate risk ratio), however, a control group would be sampled from the entire source population. Individuals in the control group would then be classified into exposed and unexposed categories. Although it may not be possible to compute the probability of a severe crash in each exposure group, the odds[4] of severe crash in each group can be computed.

The purpose of the control group is to determine the relative size of the exposed and unexposed components of the source population. The efficiency of the case-control design stems from not having to gather information on all roadway segments. Even though it precludes one from calculating the ratios of probabilities of severe crash occurrence on segments with or without a particular end treatment; it is possible to estimate ratio of odds. For rare events like severe crashes, odds ratio and risk ratio are expected to be very close to each other (19).

TABLE 2 Research objective and type of observational study design in epidemiological studies (*17*).

Objective	Common design
Prevalence	Cross sectional
Incidence	Cohort
Cause (in order of reliability)	Cohort, case-control, cross sectional
Prognosis	Cohort
Treatment effect	Controlled trial

3 In "case-control studies," two existing groups differing in outcome are identified and compared on the basis of some potential causal attribute, e.g., Severe and non-severe crashes with type of end treatment involved are often used as potential causal attributes. The units of analysis for these assessment would be crashes.

4 "Odds" are defined as the probability that the event will occur, divided by the probability that the event will not occur.

It should be noted that since Table 2 was produced by Mann (17) in the context of epidemiological studies the last row of the table refers to controlled trial as the study design for treatment effect. In the context of roadside safety equipment, a randomized control trial would be impossible to implement and the quasi-experimental design with Empirical Bayes method approach is generally used to assess treatment effect.

Safety Evaluations Through Crash Analysis

With an understanding of the basics of various study designs, the sections below include examples of studies that have applied these designs or variations of these designs for crash analyses. As part of this review the findings from the analysis have been noted where appropriate, but the focus here is on the 1.) methodological approach, 2.) evaluation question, and 3.) data collection efforts of these works. These studies are organized into two categories. In the first crashes are the unit of analysis. In the second, bothh the severity of crashes and the number of crashes (either total or severe) are modeled.

Crash Count Estimation

As part of a 2015 study, Russo (*20*) manually reviewed crash reports to identify critical supplementary information not normally available from the standard database fields for crashes involving high- tension cable median barriers. Manually extracted information included specific median cross-over category for each crash as well as which vehicle (in the case of multi-vehicle crashes) entered the median or struck the cable barrier. The manual reviews were performed to also record whether the crash involved an emergency vehicle median crossover. *From an ISPE perspective, state agencies would benefit from making such details as standard part of crash databases.* The Russo study employed EB procedure to evaluate the performance of cable median barriers by analyzing data before and after the installation. The study report also noted that experiences with cable barrier in southern states may not translate well to northern states which experience different weather characteristics and driving populations. Weather conditions, horizontal curvature, and offset of cable barrier from the roadway were found to play a role in the frequency and severity of crashes, as well as cable barrier performance. The results yielded a benefit/cost ratio for installation of cable median barriers (*20*). Candappa et al. (*21*) also used the quasi-experimental before and after study design, comparing crash frequency at each treated road section to that at untreated road sections of the same route (control sites) over the same time periods.

Martin, Mintsa-Eya, & Goubel (*22*) evaluated long-term impacts of longitudinal barriers (LBs) on motorway safety on French toll roads (*22*). The effectiveness of each type of LB was assessed as the ratio between the number of vehicles hitting the device with at least one occupant injured or killed, divided by the total number of vehicles hitting the device. A Poisson regression model combined with a robust error variance estimation procedure allowed for directly estimating relative risk measure as opposed to be being approximated by the odds ratio. The error variance estimation method used in the study was proposed by Zou (*23*) who also noted how the procedure could be implemented using SAS (one of the most widely utilized statistical analysis packages).

In 2015, Park and Abdel-Aty (*24*) noted that since the roadside elements, such as guardrail end treatments, are usually simultaneously applied to roadways and implemented at the same location, interaction effects among multiple roadside features need to be considered. In general, most previous studies have estimated single treatment effect with no attention for multiple treatments since it is difficult to consider the safety effect of any single treatment from other multiple treatments implemented at the same time using the observational before-after studies [e.g. (*25, 26*)].

To account for these implementation conditions, Park and Abdel-Aty (*24*) adopted a cross-sectional study design to develop the Crash Modification Factors (CMFs) using cross-sectional method for single and multiple treatments for different crash types and severity levels. The cross-sectional method is also known as SPFs or crash prediction models. The GLM (Generalized Linear Model) with NB (Negative Binomial) distribution model is most commonly used to develop a SPF since the function can account for over-dispersion. It was found that MARS (Multivariate adaptive regression spline) models generally provide better model fit than the GLMs and GNMs (Generalized Non-linear Models). Thus, MARS models may be better suited as part of the ISPE process for roadway locations where multiple treatments are simultaneously or near-simultaneously implemented.

Injury Severity Estimation

Through injury severity analysis it is possible to determine where there are opportunities to reduce the incidence of fatal accidents. One analysis suggests that focusing on motorcyclists and passenger vehicle side-impact crashes could reduce the number of fatal crashes by almost half (*27*). This conclusion was reached by gathering data from guardrail collisions all across the US from 2000 to 2005 and evaluating crash statistics by vehicle type, guardrail type, crash type and severity, and end result (rollover, redirect back onto the road, etc.). This assessment also showed that motorcyclists and passenger vehicle side-impact crashes (as opposed to head-on) accounted for a large portion of all fatalities in crashes involving guardrail.

Daniello and Gabler (*28*) utilized the Fatality Analysis Reporting System (FARS) database in conjunction with the General Estimates System (GES) in a 2011 study to analyze fatality risk for motorcycle crashes using data from 2004 to 2008. Three measures of crash risk were estimated, including:

- Relative fatality risks based on the most harmful event (Relative fatality risk = Fatality risk of roadside object collision/Fatality risk of ground collision) using GES database
- Relative fatality risks based on the sequence of events using FARS database which provides more detailed set of events compared to GES database.
- Distribution of most harmful event in fatal fixed object–ground crashes using FARS database

Savolainen et al. (*29*) provided a thorough review of the methodologies used for crash-injury severity analysis. These techniques range from simple, binary outcome models to more sophisticated models allowing for heterogeneous effects and correlated error terms. Because injury-severity categories are typically ordered and in sometimes closely related categories (for example no-injury and possible-injury), there may be shared unobserved effects among adjacent injury categories. The research team concluded that the appropriate methodological approach depends heavily on the available dataset, including the number of observations, quantity and quality of explanatory variables, and other data-specific characteristics.

Holdridge et al. (*16*) study analyzed the in-service performance of roadside hardware on the entire urban State Route system in the State of Washington by developing multivariate statistical models of injury severity in fixed-object crashes using discrete outcome theory. Holdridge and his research team examined the role of 'endangering factors' (driver, vehicle, environmental, and road characteristics), and their potential impact on collision severity if it involved a roadside object. Crash reports from the Washington State Highway Patrol were also reviewed to gather detailed data on every collision that occurred on the state highway system. It was concluded that components of safety devices like guardrail leading ends and bridge leading ends were associated with higher severity level while other safety devices or parts of safety device like concrete barrier, beam-guard rail faces associated with lower severity levels.

The models used in the Washington study were multivariate nested logit models of injury severity, estimated with statistical efficiency using the method of full information maximum likelihood. In other words, the entire model (i.e., Levels 2 and 3 of the nested structure shown in Figure 1) was estimated simultaneously, rather than sequentially estimating the coefficients of each level. This is known as "full information maximum likelihood" (FIML). According to Shankar et al. (30) the FIML estimator allows for consistent and statistically efficient parameter estimates. This is in contrast to the more typical sequential estimation approach, limited information maximum likelihood (LIML), which tends to be statistically inefficient and can result in biased estimates of standard errors. This study was insightful in terms of its use of *driver* injury severity in fixed-object collisions and further research was recommended to examine injury severity of *passengers* involved.

In the same study, Holdridge et al. (*16*), researchers also applied models for unordered categorical variables rather than ordered categorical variables. Even though injury severity was clearly ordered from the least-severe to the most-severe, unordered formulation allowed for greater flexibility in specification. According to Holdridge et al. (*16*):

> "Ordered models place a restriction on the effects of the explanatory factors, causing those factors to either increase the probability of greater severity or to increase the probability of lesser severity. This restriction is avoided by the use of unordered models, which means that a single explanatory factor can increase (decrease) the probability of both the greatest and least severe categories, implying a reduced (increased) probability of middle severities. This phenomenon can occur in injury severity modeling. For example, in the case of inclement weather, the number of low severity injury crashes goes up but the number of serious crashes can also go up (i.e., there is less middle ground)."

In 2014, Zou et al. (*31*) investigated 2,124 single-vehicle crashes (3,257 occupants) between 2008 and 2012 on 517 pair-matched homogeneous barrier and non-barrier segments. A binary logistic regression model with mixed effects was estimated for vehicle occupants. The segment pairing process and the use of random effects were able to handle the commonality within the same segment pair as well as the heterogeneity across segment pairs.

In another study report, Alluri et al. (*32*) compared the safety performance of G4 (1S) strong-post W-beam guardrails and cable barriers installed in the medians on freeways in Florida. The comparison was based on the percentages of barrier and median crossovers by vehicle type and crash severity. Similar to several prior research studies, data quantity became an issue when inadequate number of study sites over a short span of 1-3 years were analyzed (*33*).

Xie et al. (*34*) used Markov chain Monte Carlo algorithms to demonstrate that Bayesian models produced more reasonable parameter estimates and had improved predictive capabilities for smaller sample sizes when compared to traditional ordered probit model. For Bayesian inference, the parameters to be estimated are assumed to follow certain prior distributions. These prior distributions reflect analysts' prior knowledge about the data to be analyzed and the mean and variance for the prior estimate can be obtained based on inferences drawn other past studies.

In a 2012 published study, Xie et al. (*35*) applied a latent class logit (LCL) model that had the advantage of not restricting the coefficients of each explanatory variable in different severity functions to be the same. This made it possible to identify the impacts of the same explanatory variable on different injury outcomes. In addition, the unique model structure allowed the LCL model to better address issues pertinent to the independence from irrelevant alternatives (IIA) property. In another 2012 study, Patil et al. (*36*) demonstrated the application of weighted conditional maximum likelihood estimator in addressing the issue of underreporting of motorcycle crashes on rural Texas roads. This issue was noted as a limitation, but not explicitly addressed, by several studies.

Nearly all of safety data–related studies, those estimating injury severity and crash count models, which were reviewed were based on the use of police-reported injury data (broad discrete injury categories). More detailed data with specific injury types (head, arm, torso, etc.), lengths of hospital stays, days to recovery, may also be useful. However, such data would most likely reside in hospital databases and could be difficult to access due to privacy concerns. Savolainen et al. (*29*) pointed to the need to research based on Crash Outcomes Data Evaluation System (CODES) that links crash data with health outcome data.

Rare-Event Prediction and Analysis

Although vehicle collisions with roadside barriers occur with regularity on roads throughout the country, vehicle crashes into barrier end treatments are comparatively less frequent. Even less common are end treatment crashes that permit the conditions of the crash to be analyzed with high degrees of precision and reliability. Given the near-infinite combination of crash variables (impact location, vehicle mass, speed, angle of impact, etc.), it is rarely, if ever, possible to compare and assess the performance of end treatments that occur under controlled, experimental crash testing to that observed under field conditions. This problem is not unique to the analysis of barrier end treatments, however. In fact, this type of "rare-event" analysis is important in many fields.

Rare events are defined, most simply, as conditions that have a low probability of incidence. Because of this, it is difficult to draw meaningful and reliable conclusion about their occurrence. This can be problematic in many fields of study because many types of rare-events are also associated with high-consequence outcomes. These include, for example, catastrophic disasters associated natural hazards like earthquakes, tsunamis, hurricanes, and floods among others. Events like acts of terrorism, industrial accidents, and financial crashes are also typically considered to be "rare," but they can also have enormous societal impacts.

Because of their high potential for loss, organizations like governments, insurance companies, financial institutions, militaries, public health and disaster management agencies all have an interest in being able to forecast, model, detect, classify, and analyze rare events. As awareness of global climatological changes (among many conditions) and it's potential to significantly impact capital infrastructure and the ability to move people and goods, safely and efficiently, rare event analysis is also rapidly becoming and increasing topic of concern in transportation.

In the sections that follow, emerging knowledge from the research and application of rare-event prediction, modeling, analysis, and evaluation from fields outside of transportation are discussed. The focus of these sections is on techniques that may have application to the ISPE of barrier end treatments. And, while many of the ideas included here are theoretical, further research and experimentation may lead to ways to adapt these ideas for end treatment analysis applications.

Statistical Analysis, Evaluation, and Modeling

Wikipedia defines rare event modeling (REM) as "efforts to characterize the statistical distribution parameters, generative processes, or dynamics that govern the occurrence of statistically rare events, including but not limited to high-impact natural or human-made catastrophes. Such "modeling" may include a wide range of approaches, including, most notably, statistical models derived from historical event data and computational software models that attempt to simulate rare event processes and dynamics" (*37*). REM also encompasses efforts to forecast the occurrence of similar events over some future time horizon, which may be of interest for both scholarly and applied purposes (e.g., risk mitigation and planning).

Sampling and Interpretation Issues

Statistical analysis of rare events data—defined by King and Zeng (*38*) as "binary dependent variables with dozens to thousands of times fewer ones" was shown to be quite challenging even as such data are commonly of interest in several field beyond traffic safety and even transportation engineering. King and Zeng (*38*) also noted data collection strategies commonly used for rare event modeling may be inefficient. When one of the values of dependent variable is rare in the population, econometrics literature recommends a choice-based or endogenous stratified sampling referred as epidemiology as a case-control design (*39*).

Rare event modeling also leads to issues in evaluation of the model results. In the context of injury severity, standard measures of accuracy may not be sufficient due to rarity of most severe crash events. This issue was discussed in detail by Pande and Abdel-Aty in the context of classification of rear-end crash patterns vs. 'normal' conditions on the freeways (*40*).

An editorial article by Allison (*41*) discussed the pitfalls of using logistic regression modeling to forecast rare events. The discussion was based on the study by King and Zeng (*38*) which questioned the legitimacy of using conventional logistic regression for datasets containing rare events. It was stated that the problem was not specifically the rarity of events, but rather the possibility of a small number of cases in the rarer of the two outcomes. For example, if with a sample size of 1,000 events there are only 20 rare events as opposed to a sample size of 10,000 with 200 events, or 100,000 cases with 2,000 events, traditional statistics testing procedures will suffer from the problem of small-sample bias in the maximum likelihood estimation of the logistic model. The degree of the bias is strongly dependent on the number of cases in the less frequent of the two categories. Thus, even with a sample size of 100,000, and only 20 rare events in the sample, there is the possibility of substantial bias.

One solution to reduce the bias is Firth's penalized likelihood approach, which is a method of addressing issues of separability, small sample sizes, and bias of the parameter estimates. The method is well-known in the statistical analysis realm and can be readily implemented in statistical computing packages such as SAS' logistic regression procedure (*42*).

Data Mining Based Approaches

Data mining approaches can also help with the methodological challenges presented by traditional statistical modeling. Some of the applications where data mining methods have been applied in areas with rare yet expensive cases include, Insurance Risk Modeling (e.g. (*43*)); Hardware Fault Detection (e.g., (*44*)); and Airline No-show Prediction (e.g., (*45*)). In the subsections below, some of these data driven machine learning methods are discussed in the context of rare event modeling. Rare event analysis often needs to begin with the identification of a rare event from within a series of more routine observations. Review of research literature suggests that data mining techniques are commonly used for rare event detection. An online white paper by the Southern Methodist University College of Engineering (*46*) suggests that data mining techniques for rare events fall into two primary categories, including "supervised" and "unsupervised" techniques. Broadly, supervised techniques are classification-based methods, often using machine-learning algorithms trained using pre-classified data. In applications of these methods data is labeled as "normal" or "rare" as supervised algorithms seek to identify them with high recall or precision or both. In unsupervised detection methods, rare events are identified using techniques in which models automatically "know" what to look for without comparing to a predetermined, known conditions. In prior work this has been accomplished through spatial, temporal, or unusual variations and/or transitions between observations.

Lazarević at al. (*47*) examined the application of data mining techniques to identify and analyze rare events in security (computer network intrusion), financial (credit card fraud detection) and medical (disease diagnostics) applications. The research applied a large range of detection techniques and discussed many of the general advantages and pitfalls involved in these types of analyses. Among these methods were the creation of models of "normal" behavior to detect anomalies as deviations from this baseline. In such "outlier detection" methods, specific data points are identified because they vary from other data based on some measure.

Key performance measures and definitions were also discussed in the Lazarević work, including:

- Detection rate or "recall": ratio between the number of correctly detected rare events and the total number of all rare events;
- False alarm or "false positive" rate: ratio between the number of data records from majority class that are misclassified as rare events and the total number of data records from majority class; and
- ROC Curves: graphic representation of the trade-off between detection rate and false alarm rate.

Lazarević at al. (*47*) also pointed out (as did Pande and Abdel-Aty (*40*) in the context of freeway crashes) that classification 'accuracy' should not necessarily be regarded as a sufficient metric for evaluation because the vast majority of observations are 'normal.' This would be a key consideration in evaluating performance of the data mining based classification models employed within the ISPE process, e.g., to identify characteristics of severe crashes vis-à-vis non-severe crashes. Another overall conclusion that was reached from their work (*47*) was that the current state-of-the-art data mining techniques remain insufficient for identifying and analyzing rare events and improved designs are needed for more accurate data mining in the future.

In an earlier attempt at rare event identification, Oosterwijk et al. (*48*) applied an automated microscopy and image analysis system to detect the presence of rare fetal nucleated red blood cells. Automated techniques were used in lieu of more traditional manual screening. The results showed a benefit of over 64 percent more correct diagnosis during testing. In addition, and perhaps relevant to the ISPE processes, the author stated that another significant benefit of the technique was a reduction in personnel workloads to perform image analysis screening. Automated techniques for image analysis may be potentially useful in the ISPE process in identifying guardrails damaged in unreported crashes.

Applying another computational technique, Vannucci et al. (*49*) focused on the detection of rare patterns in unbalanced datasets from industrial manufacturing applications, particularly product defects and machine faults, using high-rate of infrequent patterns. The researchers used a combination of "internal" and "external" mathematical methods, taking advantage of the strongest aspects of both, to under- and over- sample and train the search algorithm. Another notable aspect of the research was that it looked at the issue of recognition in terms of combining the rarity of the occurrence and the importance of its detection.

In a doctoral dissertation by Huang (*50*), research focused on the identification of rare events (defined as a proportion 0.05 or less of the study sample) from a wider number of routine observations. To accomplish this, the research methodology sought to identify rare occurrences by analyzing the "right tail" of a response distribution. To test the experimental methodology, it was applied to three sets of data including one on computer CPU performance and two others on fishery bycatch. Results from these analyses were compared to the "Random Forest Method," which is a combination of multiple classification trees, and the relative advantages and disadvantages of the two methods were compared and discussed.

Forecasting and Predictive Modeling

In addition to searching for and identifying rare events, the review of research literature also showed a considerable amount of work to forecast the occurrence of rare events. Forecasting and predictive modeling techniques focused, not surprisingly, on the development and application of mathematical techniques, particularly those that could be coded into software-based applications in which significant data were available that could be searched both rapidly and with high degrees of accuracy.

A 2013 doctoral dissertation by Shi (*51*) examined the modeling, simulation design, and algorithm analysis for rare events in stochastic systems. The dissertation included an overview of these topics and a discussion and assessment of their theoretical underpinnings. In addition to presenting new and sophisticated computational theories, the author also commented on the utility and value of the various methods. Few, if any, of these, however, appeared to be directly applicable to the ISPE of guardrail end treatments.

Earlier research by Weiss and Hirsh (*52*) resulted in a genetic-based machine learning system to predict events by identifying temporal and sequential patterns in datasets, then using it to predict rare events. The concept was to predict future events based on a history of past events using enormous amounts of data. The focus of their work was on predicting hardware component failures that occur within communication networks. The process used sequences of alarms reported by monitoring software to predict atypical events. The process worked similar to the predicting fraudulent credit card purchases which are based on historical patterns of legitimate transactions.

Hoffmann et al. (*53*) developed and evaluated techniques to model and predict the occurrence of rare events. The system was based on two non-parametric techniques that make predictions as a function of discrete and continuous measurements of system variables using an extended Markov chain model and a function approximation technique based on universal basis functions (UBF). To test them, the techniques were applied to actual commercial telecommunication data that continuously measured system states. The research findings suggest that their method improved forecasting over existing approaches by an order of magnitude.

More recently, Boire (*54*) recognized that while predictive analytics can be used to identify patterns and trends that can explain future events or outcomes, these techniques perform very poorly for insurance purposes if rare events or situations are not adequately analyzed and taken into account. He suggested that in the development of insurance risk models, the models must also account for activities like fraud of major disasters; the occurrence of which may be rare and the outcomes and results are highly variable and volatile. An important point of note in this work was to make sure there are enough of rare events to both build the model then, perhaps most importantly, validate it using multiple scenarios and datasets. This points to even higher data requirements if applying data mining techniques as part of the ISPE process since they are prone to over-fitting of the training dataset if not properly validated with an independent dataset.

The forecasting of rare events is also of significant importance to governmental agencies. In a governmental request for information, the Director of National Intelligence (*55*) sought emerging methods to model and forecast rare events as well as approaches for assessing the performance of these methods, particularly those that could be generalized or adapted for a range of domains of interest to the national security agencies. While this source does not introduce any new or applicable methodologies, it included a considerable amount of key factors to consider in rare event forecast modeling. The document also defined several key terms, including a "forecast" as "a probability assigned to an event or class of events" and a "rare event" as "an event observed with very low temporal or spatial frequency relative to the parent data population or reference class, such as less than one instance in one thousand observations."

In another recent research study, Paylou et al. (*56*) discussed the potential of "penalized regression" methods to address the problem of "over-fitted" models that tend to underestimate the probability of an event in low risk patients and overestimate it in high risk patients. Because this issue can affect clinical decision making, the authors worked to develop more accurate prediction models that predicted health outcomes using risk prediction models based on patient characteristics. Among the key findings of their effort were that (*56*):

- "model overfitting could arise when the number of events is small compared with the number of predictors in the risk model;
- In an over-fitted model, the probability of an event tends to be underestimated in low risk patients and overestimated in high risk patients; and

- In datasets with few events, penalized regression methods can provide better predictions than standard regression."

The issue of overfitting being critical once again indicates that if these data-driven modeling techniques are to be used for safety evaluation within the ISPE process the analysts and agencies should ensure sufficient data is gathered not just for calibrating (or estimating) the models but also for validating the models.

Analysis, Evaluation, and Modeling

The next key area of interest was to evaluate the reliability and accuracy of identifications and predictions of rare events. While, in general, the direct application of these techniques has limited potential for the ISPE process; the following examples are helpful to show *what* these techniques have been applied to, *how* they have been applied, how aspects and the results gained from their application might be adaptable for rare highway safety event analyses.

A study sponsored by the Office of Secretary of Defense (OSD) (*57*) , requested as part of the Strategic Multi-Layer Assessment (SMA), was designed to develop a scientific/theory basis and methodology to deal combat global terrorism threats, particularly those associated with the acquisition and use of weapons of mass destruction (WMD). The application of this information would be to target areas, entities and persons of threat and then to identify and prioritize information collection requirements. Among the conclusions of the work was that *unvalidated models cannot be relied upon.* It once again empathizes the need for independent validation dataset for rare event modeling.

In an application of rare event forecast modeling to the financial markets, Bozdog et al. (58) presented a methodology to detect large price movement relative to the number of stocks that are traded, then analyzed the behavior of selected equities after the detection of such rare events. In the paper, the researchers provided methods to calibrate rules based on event detection and apply the technique to thousands of equity trades over a five day period. Among the findings of this work were that 1.) it took a larger volume of occurrences to observe a rare event for a highly-traded equity stocks than for those that are less traded and 2.) the price recovery after a rare event is much faster for highly traded stocks than for less liquid stocks. Based on this, the authors concluded that the method would be useful to identify irregular (perhaps fraudulent/illegal) trades.

Research by Görg et al. (*59*) used discrete event simulation as a method of forecasting rare events in communication networks. Their method was proposed because prior simulation methods required lengthy run times, of perhaps even months and years. The new method, named RESTART/LRE, provided a multi-step simulation approach that reduced simulation run time by several orders of magnitude (from years to just minutes) using "importance splitting," in which system states that lead to the rare event were saved and used as the starting point for new simulation sub-runs.

A study by Cai et al. (*60*) applied meta-analysis to combine data from multiple various clinical datasets and study results to evaluate treatment efficacy. The authors recognized that since adverse events are typically rare, standard methods including fixed-effect assumption to combine results from multiple studies, do not work well in this setting. Where many common methods of rare event analysis use fixed or random effects models to combine effect estimates obtained separately in individual studies; in this work the authors proposed and tested an alternative approach based on Poisson random effects models to make inferences about the relative risk between two treatment groups. The results of the simulated application of their method showed that it performed well when the underlying event rates were low. This approach to rare event modeling may be useful in combining results from past research or multiple ongoing research efforts for evaluations within an ISPE framework.

Classification and Datasets

A key need in the analysis of rare events are databases containing observation records of "events" and "non-events." Depending upon the field of study; ease of collection and storage; the frequency of occurrence of the rare events, among many other variables; these data records and the ease of their collection, storage, and retrieval; may or may not be large, expensive, and complicated to both collect and store. The review of literature related to the classification and storage of rare event data showed a large amount of activity in two fields related to these issues.

The first were those where activity records could be accessed, reviewed, and analyzed rapidly and cheaply. These include computer network transactions, communication systems, and manufacturing processes where thousands, millions, and more observations can be made within a short period of time then retrieved for study within a matter of seconds. The second group includes activity domains where the occurrence of rare events may be considerably more difficult, time consuming and costly to acquire, but where such occurrences, when they do happen, are likely to result in costly outcomes and highly adverse consequences, particularly those related to life safety and economics. These low probability/high consequence events include terrorist attacks, extreme weather events, earthquakes, aviation accidents, and the like.

To illustrate a wide variety of existing databases used for rare event analysis, a few of them are noted below. These databases are also helpful to illustrate the breadth of fields in which rare event analysis is being conducted.

- **Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat)** (online: *http://earthquake.usgs.gov/earthquakes/search/*)
- **Aviation Safety Database** (online: *http://aviation-safety.net/database/*)
- **Database of Radiological Incidents and Related Events** (online: *http://www.johnstonsarchive.net/nuclear/radevents/*)
- **Global Health Atlas** (online: *http://apps.who.int/globalatlas/default.asp*)
- **Global Volcanism Program** (online: *http://www.volcano.si.edu/search_eruption.cfm*)
- **NOAA Natural Hazards** (online: *http://www.ngdc.noaa.gov/hazard*
- **U.S. National Flood Insurance Program** (online: *http://www.fema.gov/policy-claim-statistics-flood-insurance/policy-claim-statistics-flood-insurance/policy-claim-13*)

While these datasets may not be directly applicable to the ISPE program, given the rapid rates at which data gathering capabilities are expanding, these sets may be useful tool in training agency personnel in rare event modeling.

Rare-Event Modeling: Findings and Application Potential

The review of rare event identification, assessment, modeling, forecasting, and classification revealed many of the motivations, needs, practices, and applications associated with these analyses. Broadly speaking, the literature showed that the advancements were concentrated in fields in which rare events have high consequences, most notably related to life health/safety and economic gain/loss. Considerable research and application activity was also observed in fields where rare events can be observed from large databases that can be searched relatively simply, rapidly, and inexpensively with high accuracy. Because of this, these activities tend to utilize automated, computationally-oriented, digital processing methods that facilitate the evaluation of individual observations from within very large data sets, particularly those in which data is produced in very high volume and at very high rates - like computerized and mass manufacturing processes.

Based on these findings majority of methods utilized by the rare event modeling studies reviewed herein *do not* show potential for direct application to the ISPE of barrier end treatments because they require large amount of training and yet more independent validation dataset. This is not to suggest, however, that nothing can be learned from these findings. While the review did not show how to carry out such analyses, specifically, it does suggest considerations and ideas that can be taken into account when undertaking rare event analyses. Of particular note are aspects of prior research and application that maybe be modified and adapted for ISPE use as well as an understanding of the difficulties and limitations of its use. Also, there are two methods that should be applicable in safety evaluations carried out as part of the ISPE process. One is the Firth's penalized likelihood approach when modeling rare events using logistic regression and the other is the meta-analysis approach used by Cai et al. (*60*). The later could be useful in combining results from evaluations carried out at multiple sites.

The review also suggests a number of fundamentally related conditions of rare event analysis that could be worthy of consideration, more broadly. The most obvious of these is that the low frequency of rare events makes modeling and forecasting inherently difficult. This is particularly true from the standpoint of maintaining statistical validity. Given the inevitable low number of observations in a data set, the literature suggests that conventional statistical inference techniques can be problematic, if not impossible, to utilize (*55*). The prior studies also show that an understanding of the statistical assumptions of the underlying models is a key consideration. For example, it is important to know if a predictive model is based on a specific heavy-tailed distribution relating probability to frequency and whether the magnitude of the modeled distribution is empirical or theoretical. The "limits" and boundary conditions of a model are also an important considerations. These include the time horizon of forecasts, data requirements, availability of outcome/evaluation data, resource requirements, or other limiting factors like study time and cost and the labor effort required to carry it out. Other key considerations include required aspects of the modeling approaches such as the processes used to validate or assess the accuracy and/or usefulness of a forecast and the strengths and limitations of them.

CRITICAL ASSESSMENT OF METHODS

In this section, the relevant methods/study designs reviewed in this study are comparatively presented in a two-dimensional space defined by i.) the objective of evaluation on the horizontal axis (See bottom of Table 3) and ii.) cost/effort involved in collection of the data required by the approach on the vertical axis (Higher cost/effort as one moves up the table). It should be noted that unlike the vertical axis, categorization on the horizontal axis of the table does not necessarily imply any ordered levels and simply depends on analyst/agency objectives.

At the low-cost end of the table are methods for preliminary identification of safety issue (Column 1; Table 3). The cost of gathering data for these methods is low as this can be accomplished using existing nationwide databases. For example, preliminary risk identification for locations with and without guardrail may be accomplished using existing federal level datasets (e.g., FARS and GES).

TABLE 3 Data collection Cost/Effort and Evaluation Objective for Study Designs Reviewed in the Paper

Cost/effort of Gathering Information/Data	Estimation Relative Fatality Risk (e.g., Daniello and Gabler (28))	Cross-sectional study design (e.g., Park and Abdel-Aty (24))	Quasi-experimental design (Empirical Bayes)	Traditional Multinomial logit and ordered Probit Models (34) Logit Model enhanced with Bayesian Priors (34)	Examine failure causes by examining damages and mechanical performance of the end treatments (4).
	Preliminary Identification of Safety Issue	**Estimation of CMFs for guardrail end treatments**	**Compare effectiveness of end-treatment options (based on changes in crash experience)**	**Compare effectiveness of end-treatment options (based on severity outcome)[5]**	**Provide information to support future end treatment designs**
	Evaluation Objective(s)				

[5] Case-control design would be an appropriate choice of study design for this evaluation objective.

The studies that involve next level of data collection costs are cross-sectional study design. Crash Modification Factors (CMFs) can be estimated using cross-sectional study design by gathering crash data from several locations with as well as without a specific treatment. Data used in cross-sectional study (i.e., development of an SPF) will typically be supplied by existing state level crash and equipment inventory databases.

Specific evaluation of devices, however, may require examining change in crash experience of location before and after installation of the end treatments (Column 3) and/or modeling of severity outcomes (Column 4). Some such studies may be accomplished using existing state level crash and equipment inventory databases. However, it should also be recognized that manual review of crash reports will typically enhance Before/After studies examining crash experience of specific sites [e.g., (*20*)]. Similarly, severity outcome studies may be enhanced by linking Crash Outcomes Data Evaluation System (CODES) that links crash data with health outcome data (*29*).

Among the severity outcome studies, traditional model formulations, e.g., multinomial logit or ordered probit model require large samples of data. However, these formulation could potentially be modified with Bayesian priors based on experience and past research to work with smaller sample sizes, thereby reducing the cost of data collection (Hence, lower placement of these methods in Table 3). Meta-analysis method proposed by Cai (*60*) in the context of epidemiological studies may similarly allow for utilizing results from past studies to reduce sample size requirements.

Lastly, if evaluation objective is to improve future safety equipment design and/or improve crash tests to better reflect in-service conditions of the devices; that would require significant cost of data collection, including an On-Call Investigation Team to assess the damage before the end treatments are repaired. Hence, these types of studies are placed on the high end of the cost spectrum (See the top right corner of Table 3).

CONCLUSIONS

The review of ISPE processes and relevant evaluation methodologies for guardrail end treatment and associated roadway safety-related devices was conducted over three primary categories of research and applications. These included:

1. actual ISPE practice;
2. the application of crash/safety data and statistical analysis to inform and enhance ISPE process; and
3. techniques from other fields of study and practice to assess and analyze rare events which may be adapted for use in ISPE.

Table 4 summarizes the findings as they relate to the ISPE process goals outlined at the outset of this report.

TABLE 4 Methods Appropriate for ISPE Process Objective Specified by the Guardrail Committee

ISPE Process Objective	Applicable Methods and Databases
1. Developing data capture protocols for in-service performance evaluation	• Review findings from ISPE process for Texas (3) and Arizona (4). • Standardize collection of data elements manually collected by Russo (20) for road departure crashes.
2. Surveying methods for analyzing engineering rare events	• Review findings from this report
3. Comparing the effectiveness of various guardrail end treatment designs in minimizing injury severity in a crash,	• Nested severity outcome model estimation through existing crash, roadway characteristics, and roadside equipment database
4. Evaluating the impact of installation and maintenance practices and environmental conditions on in-service safety performance	• Quasi-experimental Before/After analysis through EB procedure through existing crash, roadway characteristics, and roadside equipment database
5. Estimating benefits and costs of a program of replacing installed end treatments with end treatments that perform better;	• Quasi-experimental Before/After analysis through EB procedure followed by Benefit/Cost Analysis
6. Providing data to support the improvement of end treatment designs;	• Analyze mechanics of each collision involving guardrail end treatment. • Continuous monitoring of roadside equipment through intelligent sensors
7. Refining crash test procedures such that it better reflect the in-service performance	• Analyze mechanics of each collision involving guardrail end treatment.

The review findings showed that while some states have experience in conducting guardrail and end section ISPEs and may be acquainted with some of the findings of this study, many – if not most – will not. Below the authors summarize several notable findings that could be of value to transportation agencies tasked with conducting ISPE of safety devices in their inventory. Although not all of the practices will be applicable for all agencies, it is anticipated that many of the unique and innovative effective practices as well as lessons learned from prior work could be appropriate for incorporation into future practice guidance. This could occur directly or, more likely through the modification and adaptation of these practices to fit the location specific goals of an agency.

From the perspective of current ISPE practice, the review showed that there has been an extensive and varied history of activity in the US and Canada, led primarily by State

Departments of Transportation. Practices vary from extensive and systematical practices, to quite informal activities. Geographic difference were also evident. This finding was also consistent with a related prior study from Texas (*3*) that noted:

> "Each State DOT differs in terms of available and maintained information sources, procedures within the road safety management process, organizational structure, and characteristics (e.g., size, geographic location, etc.). Although there is certainly common ground in terms of the objectives of the ISPE process, it is not possible to have a "one size fits all" methodology for all state departments of transportation."

Practices ranged from those developed by the Arizona Department of Transportation (ADOT), to a virtual absence of documented efforts in other states. The ADOT process is noteworthy because incorporates a very specific and prescriptive, four-level process that links data collection from a variety of sources and personnel dedicated to the review and analysis of road safety hardware.

Considerable efforts to conduct performance reviews have also taken place at a national level using data collected over multiple states. Description of results of these studies as well as the sources and methods employed to collect and analyze data was discussed earlier in this report. As illustrated in Table 3 methodology used as well as the cost, time, and effort to collect data to implement the methodology is closely dependent on the analyst's objective. This broad-based review also revealed other more subtle nuances that should be taken into account in future work. For example, a 2011 published study (*28*) suggested that crash tests for end-treatment designs do not adequately account for motorcycle related crashes. Hence, collisions involving motorcycles may need to be monitored more closely during the ISPE process to identify potential safety issues.

The review of methodologies that incorporate highway crash data also showed the breadth and variation of work that has been undertaken in this area of focus. It was clear that safety-related databases of crash characteristics for vehicle collisions into end treatments are a valuable input to the ISPE process. For example, such data can be used to determine how similar or dissimilar crash conditions may be to the standardized test conditions outlined in MASH. Similar to the findings of ISPE field practices, the review also found that a wide range of techniques and data sources have been employed to analyze the in service performance of road safety features. Of particular note is the range of statistical data acquisition techniques, databases, and analysis techniques that have been explored in prior research.

A 2014 study from Korea (*61*) mentioned context-awareness system based on acceleration sensors for guardrails that may be used to continuously monitor the guardrail while in-service. Advances in the sophistication of such devices, coupled continued decreases in their costs, could make it possible to apply them in permanent field installations to monitor the evolving condition of roadside equipment as well to record a more accurate number and characteristics of collisions involving various devices. In a related finding of note, the future incorporation of medical and injury data may be useful in analytical severity outcome studies. While such statistics could be a reflection of vehicle design, use of occupant restraint systems, emergency response capability, and other medical-related issues, the inclusion of such data is nonetheless useful to gauge the performance of roadside safety systems in terms of more precise human injury costs.

The findings from the final area of review, rare event identification, assessment, modeling, forecasting, and classification; suggested that work in other fields generally shows limited potential for application in ISPE of guardrail end treatments. It was apparent that a considerable amount of research and application has taken place in fields for which rare events have high consequences. Activity was also apparent in fields from which rare events could be observed from large databases that can be searched relatively simply, rapidly, and inexpensively with high accuracy using automated, computationally-oriented, digital processing, particularly those in which data is produced in very high volume and at very high rates - like computerized and mass manufacturing processes. However, this type of activity differs fairly significantly from road safety devices where data is not nearly as plentiful or easy to collect or autonomously analyze.

Other associated work from rare event analysis was also somewhat similar to prior work in transportation, particularly in terms of the cost-benefit trade-offs between cost, effort, complexity and the utility of the results. Although data mining based non-parametric approaches for rare-event modeling could have some potential, they like most of the other techniques did not clearly suggest high potential for direct application in the ISPE of guardrail end treatments. However, as agency capabilities to continuously monitor infrastructure and gather and archive relevant safety data advances, it is possible that some of these methods may have increased application potential.

REFERENCES

1. "Fatality Facts." [Online]. Available: http://www.iihs.org/iihs/topics/t/roadway-and-environment/fatalityfacts/fixed-object-crashes. [Accessed: 24-Jun-2016].
2. A. A. of S. Highway and T. O. T. F. for R. Safety, *Roadside design guide*. AASHTO, 2011.
3. I. van Schalkwyk, R. P. Bligh, D. C. Alberson, D. L. Bullard Jr, D. Lord, and S.-P. Miaou, "Developing an In-Service Performance Evaluation (ISPE) for Roadside Safety Features in Texas," FHWA/TX-05/0-4366-1, 2004.
4. K. Mak and D. Sicking, "Continuous evaluation of in-service highway safety feature performance," Arizona Department of Transportation, Final Report 472, 2002.
5. M. H. Ray, J. Weir, and J. Hopp, "In-Service Performance of Traffic Barriers. NCHRP Report 490," *Transp Res Board Wash. DC*, 2003.
6. D. L. Sicking, K. K. Mak, J. R. Rohde, and J. D. Reid, "Manual for assessing safety hardware," *Wash. DC Am. Assoc. State Highw. Transp. Off.*, 2009.
7. "In-Service Performance of Energy-absorbing W-beam Guardrail End Treatments: Phase 1 | Policy Studies." [Online]. Available: http://www.trb.org/PolicyStudies/WbeamGuardrailTreatments.aspx. [Accessed: 24-Jun-2016].
8. M. H. Ray, C. Silvestri, C. E. Conron, and M. Mongiardini, "Experience with cable median barriers in the United States: Design standards, policies, and performance," *J. Transp. Eng.*, vol. 135, no. 10, pp. 711–720, 2009.
9. H. C. Gabler, D. J. Gabauer, and W. L. Szalaj, "Safety Audit of Fatalities and Injuries Involving Guide Rail," FHWA-NJ-2007-001, 2006.
10. "Safety Analysis of Extruding W-Beam Guardrail Terminal Crashes | FHWA Review of ET-Plus | Federal Highway Administration." [Online]. Available: http://www.fhwa.dot.gov/guardrailsafety/safetyanalysis/. [Accessed: 27-Jun-2016].
11. K. R. Agent, "Evaluation of the ET2000 Guardrail End Treatment," 2004.
12. R. W. Esligar and E. D. Hildebrand, "Crash Performance of Energy-Absorbing Guide Rail Terminals," in *2012 CONFERENCE AND EXHIBITION OF THE TRANSPORTATION*

ASSOCIATION OF CANADA-TRANSPORTATION: INNOVATIONS AND OPPORTUNITIES, 2012.

13. "Guidelines_for_Traffic_Barrier.pdf." [Online]. Available: https://www.roads.maryland.gov/ohd/Guidelines_for_Traffic_Barrier.pdf. [Accessed: 11-Sep-2016].
14. "Pub652.pdf." [Online]. Available: https://www.dot.state.pa.us/public/pubsforms/Publications/Pub%20652.pdf. [Accessed: 11-Sep-2016].
15. E. Hauer, D. W. Harwood, F. M. Council, and M. S. Griffith, "Estimating safety by the empirical Bayes method: a tutorial," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1784, pp. 126–131, 2002.
16. J. M. Holdridge, V. N. Shankar, and G. F. Ulfarsson, "The crash severity impacts of fixed roadside objects," *J. Safety Res.*, vol. 36, no. 2, pp. 139–147, 2005.
17. C. J. Mann, "Observational research methods. Research design II: cohort, cross sectional, and case-control studies," *Emerg. Med. J.*, vol. 20, no. 1, pp. 54–60, 2003.
18. American Association of State Transportation Officials, *Highway safety manual*, 1st ed., vol. 1, 3 vols. American Association of State Highway and Transportation Officials, 2010.
19. J. Kenneth, *Rothman. Epidemiology: an introduction*. New York: Oxford University Press, 2002.
20. B. J. Russo, "Guidance on the application of cable median barrier: tradeoffs between crash frequency, crash severity, and agency costs," Iowa State University, 2015.
21. N. Candappa, A. D'Elia, B. Corben, and S. Newstead, "Wire rope barrier effectiveness on Victorian roads," in *Proceedings of the Australasian road safety research, policing and education conference*, 2011, vol. 15.
22. J.-L. Martin, C. Mintsa-Eya, and C. Goubel, "Long-term analysis of the impact of longitudinal barriers on motorway safety," *Accid. Anal. Prev.*, vol. 59, pp. 443–451, 2013.
23. G. Zou, "A modified poisson regression approach to prospective studies with binary data," *Am. J. Epidemiol.*, vol. 159, no. 7, pp. 702–706, 2004.
24. J. Park and M. Abdel-Aty, "Assessing the safety effects of multiple roadside treatments using parametric and nonparametric approaches," *Accid. Anal. Prev.*, vol. 83, pp. 203–213, 2015.
25. D. L. Harkey, N. C. H. R. Program, A. A. of S. Highway, and T. Officials, *Accident modification factors for traffic engineering and ITS improvements*, vol. 617. Transportation Research Board, 2008.
26. N. Stamatiadis, D. Lord, J. Pigman, J. Sacksteder, and W. Ruff, "Safety impacts of design element trade-offs for multilane rural highways," *J. Transp. Eng.*, vol. 137, no. 5, pp. 333–340, 2010.
27. H. C. Gabler and D. J. Gabauer, "Opportunities for reduction of fatalities in vehicle-guardrail collisions," in *Annu Proc Assoc Adv Automot Med*, 2007, vol. 51, pp. 31–48.
28. A. Daniello and H. Gabler, "Effect of barrier type on injury severity in motorcycle-to-barrier collisions in North Carolina, Texas, and New Jersey," *Transp. Res. Rec. J. Transp. Res. Board*, no. 2262, pp. 144–151, 2011.
29. P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives," *Accid. Anal. Prev.*, vol. 43, no. 5, pp. 1666–1676, 2011.
30. V. Shankar, F. Mannering, and W. Barfield, "Statistical analysis of accident severity on rural freeways," *Accid. Anal. Prev.*, vol. 28, no. 3, pp. 391–401, 1996.
31. Y. Zou, A. P. Tarko, E. Chen, and M. A. Romero, "Effectiveness of cable barriers, guardrails, and concrete barrier walls in reducing the risk of injury," *Accid. Anal. Prev.*, vol. 72, pp. 55–65, 2014.
32. P. Alluri, A. Gan, K. Haleem, and J. Mauthner, "Safety performance of g4 (1s) w-beam guardrails versus cable median barriers on Florida's freeways," *J. Transp. Saf. Secur.*, vol. 7, no. 3, pp. 208–227, 2015.

33. P. Alluri, K. Haleem, and A. Gan, "In-service Performance Evaluation of Cable Median Barriers 2 on Florida's Limited Access Facilities 3," in *Transportation Research Board annual meeting, 92nd, 2013, Washington, DC*, 2013.
34. Y. Xie, Y. Zhang, and F. Liang, "Crash injury severity analysis using Bayesian ordered probit models," *J. Transp. Eng.*, vol. 135, no. 1, pp. 18–25, 2009.
35. Y. Xie, K. Zhao, and N. Huynh, "Analysis of driver injury severity in rural single-vehicle crashes," *Accid. Anal. Prev.*, vol. 47, pp. 36–44, 2012.
36. S. Patil, S. R. Geedipally, and D. Lord, "Analysis of crash severities using nested logit model—accounting for the underreporting of crashes," *Accid. Anal. Prev.*, vol. 45, pp. 646–653, 2012.
37. "Rare events," *Wikipedia, the free encyclopedia*. 19-May-2016.
38. G. King and L. Zeng, "Logistic regression in rare events data," *Polit. Anal.*, vol. 9, no. 2, pp. 137–163, 2001.
39. N. E. Breslow, "Statistics in epidemiology: the case-control study," *J. Am. Stat. Assoc.*, vol. 91, no. 433, pp. 14–28, 1996.
40. A. Pande and M. Abdel-Aty, "Comprehensive analysis of the relationship between real-time traffic surveillance data and rear-end crashes on freeways," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1953, pp. 31–40, 2006.
41. "Logistic Regression for Rare Events | Statistical Horizons." [Online]. Available: http://statisticalhorizons.com/logistic-regression-for-rare-events. [Accessed: 29-Jun-2016].
42. SAS Institute, *Base SAS 9.4 Procedures Guide*. SAS Institute, 2014.
43. E. P. Pednault, B. K. Rosen, and C. Apte, *Handling imbalanced data sets in insurance risk modeling*. IBM TJ Watson Research Center Yorktown Heights, NY, USA, 2000.
44. C. Apte, S. Weiss, and G. Grout, "Predicting defects in disk drive manufacturing: A case study in high-dimensional classification," in *Artificial Intelligence for Applications, 1993. Proceedings., Ninth Conference on*, 1993, pp. 212–218.
45. R. D. Lawrence, S. J. Hong, and J. Cherrier, "Passenger-based predictive modeling of airline no-show rates," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2003, pp. 397–406.
46. "Rare Event." [Online]. Available: http://lyle.smu.edu/cse/dbgroup/rare.html. [Accessed: 28-Jun-2016].
47. A. Lazarević, J. Srivastava, and V. Kumar, *Data mining for analysis of rare events: A case study in security, financial and medical applications*. 2004.
48. J. C. Oosterwijk, C. F. Knepflé, W. E. Mesker, H. Vrolijk, W. C. Sloos, H. Pattenier, I. Ravkin, G.-J. B. van Ommen, H. H. Kanhai, and H. J. Tanke, "Strategies for rare-event detection: an approach for automated fetal cell detection in maternal blood," *Am. J. Hum. Genet.*, vol. 63, no. 6, pp. 1783–1792, 1998.
49. M. Vannucci, V. Colla, G. Corbo, and S. Fera, "Detection of rare events within industrial datasets by means of data resampling and specific algorithms," *Int. J. Simul. Syst. Sci. Technol.*, vol. 11, no. 3, pp. 1–11, 2010.
50. W. Huang, "Methods to Extract Rare Events," University of California, Los Angeles.
51. Y. Shi, *Rare events in stochastic systems: modeling, simulation design and algorithm analysis*. Columbia University, 2013.
52. G. M. Weiss and H. Hirsh, "Learning to predict extremely rare events," in *AAAI workshop on learning from imbalanced data sets*, 2000, pp. 64–68.
53. G. A. Hoffmann, F. Salfner, and M. Malek, "Advanced failure prediction in complex software systems," 2011.
54. R. Boire, "Predicting Rare Events In Insurance," *Predictive Analytics Times*, 12-May-2015. .
55. "Forecasting Rare Events." [Online]. Available: https://www.iarpa.gov/index.php/working-with-iarpa/requests-for-information/forecasting-rare-events. [Accessed: 28-Jun-2016].
56. M. Pavlou, G. Ambler, S. R. Seaman, O. Guttmann, P. Elliott, M. King, and R. Z. Omar, "How to develop a more accurate risk prediction model when there are few events," 2015.

57. D. McMorrow, "Rare Events," Oct. 2009.
58. D. Bozdog, I. Florescu, K. Khashanah, and J. Wang, "Rare Events Analysis for High-Frequency Equity Data," *Wilmott*, vol. 2011, no. 54, pp. 74–81, 2011.
59. C. Görg, E. Lamers, O. Fu\s s, and P. Heegaard, "Rare event simulation," in *Modeling and Simulation Environment for Satellite and Terrestrial Communications Networks*, Springer, 2002, pp. 365–396.
60. T. Cai, L. Parast, and L. Ryan, "Meta-analysis for rare events," *Stat. Med.*, vol. 29, no. 20, pp. 2078–2089, 2010.
61. S.-W. Jang, S.-Y. Cho, and G.-S. Lee, "An intelligent guardrail context-awareness system based on acceleration sensors in ubiquitous sensor networks," *Int. J. Distrib. Sens. Netw.*, vol. 2014, 2014.