TRANSPORTATION RESEARCH

# RECORD

JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

NO. 1784

# Statistical Methodology

## Applications to Design, Data Analysis, and Evaluation

Safety and Human Performance

TRANSPORTATION RESEARCH BOARD
*OF THE NATIONAL ACADEMIES*

# Diagnostic Methodology for the Detection of Safety Problems at Intersections

Jake Kononov and Bruce N. Janson

**There is an established consensus among traffic safety researchers that a nonlinear relationship exists between traffic exposure and safety. This relationship is reflected by the safety performance functions (SPFs) calibrated for various classes of roads and intersections. One of the main uses of SPFs is to identify locations with potential for accident reduction. While this application is certainly important, the use of SPFs provides no information related to the nature of the accident occurrence. Without being able to relate accident frequency and severity to roadway geometrics, traffic control devices, roadside features, roadway condition, driver behavior, or vehicle type, it is not possible to develop effective countermeasures. A methodology was developed to provide guidance in diagnostics of safety problems, recognition of accident patterns, and development of appropriate countermeasures. Considering that traffic accidents can be viewed as random Bernoulli trials, it is possible to detect deviation from the statistical process by computing observed cumulative probability for each of the accident characteristics. Detection of an accident pattern at an intersection suggests the presence of an element in the roadway environment that triggered a deviation from a random statistical process in the direction of reduced safety. Identification of such an element always provides a critical clue to accident causality.**

There is an established consensus among traffic safety researchers that a nonlinear relationship exists between traffic exposure and safety. This relationship is reflected by the safety performance functions (SPFs) calibrated for various classes of roads and intersections. One of the main uses of SPFs is to identify locations with potential for accident reduction. While this application is certainly important, its use is limited to identifying sites that exhibit accident frequency higher than expected for a specific facility at a specific level of average daily traffic (ADT). SPFs provide no information, however, related to the nature of the accident occurrence; they only speak to the magnitude of the problem. Without being able to relate properly and systematically accident frequency and severity to roadway geometrics, traffic control devices, roadside features, roadway condition, driver behavior, or vehicle type, it is not possible to develop effective countermeasures. In other words, there can be no effective treatment without accurate diagnosis.

In the field of medicine, physicians are expected to spend a minimum of 3 years in apprenticeship after graduation from medical school. During the periods of internship and residency, physicians are learning how to recognize diseases as well as how to treat them. In contrast to medicine, transportation engineers are like doctors who are trained only how to administer treatment without learning the science of diagnostics. There is no established course of instruction at the graduate level of civil engineering curriculum that would provide a definitive methodology on how to relate accident causality to the roadway environment. There is also very little reliable information in the research literature on the subject. Most research efforts are focused on the development of accident prediction models and identification of “black spots.” It is somehow implied that transportation engineering professionals will always know how to treat a high accident location once it has been identified, when in reality very little is known on the subject.

In the course of the study of safety assessments of hundreds of intersections, a new methodology was developed to provide guidance in diagnostics of safety problems and developing appropriate countermeasures. The data set was compiled using accident and traffic data for different classes of intersections over a period of 8 years. A framework of normative parameters was developed to provide guidance in the diagnostics of accident causality and recognition of accident patterns. Considering that traffic accidents can be viewed as random Bernoulli trials, it is possible to detect deviation from the statistical process by computing observed cumulative probability for each of the accident characteristics. Detection of an accident pattern at an intersection suggests a presence of an element or elements in the roadway environment that triggered a deviation from a random statistical process in the direction of reduced safety. Identification of such an element or elements always provides a critical clue to accident causality. The diagnostics process of highway safety problems on a section of road is in many ways similar to making a medical diagnosis. While diagnostics is an integral part of medicine, much remains to be done by the transportation engineering profession in order to institutionalize this important component of highway safety.

## REVIEW OF EXTANT LITERATURE

Over the last 50 years of modern road building it was somehow implied that transportation engineering professionals will always know how to diagnose the nature of accident problems at a high accident location once it has been identified, when in reality very little is known on the subject. This state of affairs is best expressed by Hauer (*1*):

> If the site has been identified because its accident record is unusual, one has also to find out why. Thus, the detailed safety analysis stage is akin to a process of medical diagnosis, with perhaps a keener awareness of costs and budgets, a process requiring knowledge of causes, effects, and economics. One might expect that this task would be performed by specialists whose training in this matter is extensive and based on knowledge of fact. Unfortunately, this is not so. For some reason, perhaps because of a fascination with matters statistical or perhaps because it is a headquarters function, a great deal of thought has been

J. Kononov, Colorado Department of Transportation, Region 6, 2000 South Holly, Denver, CO 80222. B. N. Janson, Department of Civil Engineering, University of Colorado at Denver, 1200 Larimer Street, Denver, CO 80217.

> devoted to the identification stage. Much less has been written about, or taught to engineers, how to conduct a detailed safety analysis of a site. Yet, not common sense, practical experience, engineering judgement, or the usual highway and traffic engineering lore is a sufficient guide. To be effective, it is not enough to produce reasonable lists of candidate sites to be investigated in the order of priority. It is also necessary to equip the engineer with the training and the tools to make a safety diagnosis on the basis of the specific kinds of accidents that have occurred, the conditions in which they occurred, and the characteristics of the site. Furthermore, it is necessary to give the engineer realistic estimates of what safety improvements can be expected. This, at present, is a tall order.

Once again the field of medicine can provide conceptual and methodological guidance on how to formulate a solution. In the United States, the initial impetus for developing a classification of mental disorders was the need to collect statistical information. In 1917, the Committee on Statistics of the American Psychiatric Association (APA), together with the National Commission on Mental Hygiene, formulated a plan that was adopted by the Bureau of the Census for gathering uniform statistics across mental hospitals.

In 1952 APA (*2*) developed an authoritative guide on diagnostics of mental disorders entitled *Diagnostic and Statistical Manual of Mental Disorders* (DSM). In part because of the lack of widespread acceptance of the mental disorder diagnostic categories, the World Health Organization (WHO) sponsored a comprehensive review of diagnostic issues that was conducted by the British psychiatrist Erwin Stengel. According to APA (*3*), "His report can be credited with having inspired many of the recent advances in diagnostic methodology—most especially the need for explicit definitions as a means of promoting reliable clinical diagnosis." Since 1952 there have been three more editions of DSM. Over the last half-century, thousands of psychiatrists systematically collected data to advance diagnostic methodology. In contrast to medicine, no such undertaking by the transportation engineering profession has taken place. In highway safety, just as in medicine, there can be no effective treatment without accurate diagnosis.

In the fourth edition of DSM (DSM-IV) (*3*), APA cautions that

> [t]he diagnostic categories, criteria, and textual descriptions are meant to be employed by individuals with appropriate clinical training and experience in diagnosis. It is important that DSM-IV not be applied mechanically by untrained individuals. The specific diagnostic criteria included in DSM-IV are meant to serve as guidelines to be informed by clinical judgement and are not meant to be used in a cookbook fashion.

Furthermore, APA states that "The proper use of these criteria requires specialized clinical training that provides both a body of knowledge and clinical skills." A similar caution is relevant to the practice of transportation engineering, yet at present a very limited factual knowledge base exists to assist transportation professionals in making diagnostic decisions on safety problems or to provide effective training in this area.

## METHODOLOGY

Overrepresentation in the number of accidents above the expected or normal threshold predicted by the SPF is only one of many indicators of a potential for accident reduction (and it appears that it may not be the best one). Accident type, severity, road condition, spatial distribution of accidents, and lighting conditions are only a few of the many important symptoms of the accident problem. Furthermore, in many cases, factors other then overrepresentation in frequency are better predictors of susceptibility to corrective countermeasures.

It is difficult to determine a specific form for the distribution of accidents; therefore, the problem lends itself well to a nonparametric approach, which does not require assumptions about the shape of the underlying distribution. Accident occurrence as a process can be thought of as a sequence of Bernoulli trials where the following holds true:

- There are only two outcomes at each trial or observation: an accident of a specific type has or has not occurred;
- The probability of "success" is the same for each trial: the probability of occurrence of a specific accident-related event (e.g., overturning) is the same every time an accident has occurred;
- The trials are independent: each accident is completely independent from the previous or the following one; and
- There are a finite number of trials.

The following terminology can be adapted to provide the analytical framework of the pattern recognition through direct diagnostics of accident distribution profile.

- $SF_i$—denotes a specific SPF representing roadway segment or an intersection;
- $\boldsymbol{X_{ai}}$ $[\boldsymbol{X_{a1}}, \boldsymbol{X_{a2}}, \ldots, \boldsymbol{X_{an}}]$—represents a feature vector comprised of accident listing of the roadway segment directionally arranged in relation to the roadway reference system, or reflecting an accident listing at an intersection;
- $P(SF_i)$—the probability that we are presented with a SPF $SF_i$;
- $P(N_{ai} / SF_i)$—the probability that $N_{ai}$ accidents of specific type would be observed given a SPF $SF_i$;
- $P_i$—Bernoulli probability of observing a specific crash-related characteristic during each accident event (developed as an average percentage for various intersection types over a period of 8 years); and
- $P(SF_i / N_{ai})$—the conditional probability that we are presented with a SPF $SF_i$ given a feature vector $\boldsymbol{X_{ai}}$, containing $N_{ai}$ accidents of specific type.

Assume that feature vector $\boldsymbol{X_{ai}}$ represents a sample of accident history drawn from a roadway facility represented by a SPF $SF_i$. The probability that exactly $N_{ai}$ accidents of a specific type will be observed out of the total of $N_{ti}$ accidents is given by the binomial distribution:

$$\boldsymbol{X_{ai}} \in SF_i \therefore P(N_{ai}, N_{ti}, P_i) = \binom{N_{ti}}{N_{ai}} P_i^{N_{ai}} (1 - P_i)^{N_{ti} - N_{ai}} = P(N_{ai} / SF_i)$$

where $N_{ai} = 0, 1, 2, \ldots, n$ accidents, and

$$\binom{N_{ti}}{N_{ai}} = \frac{N_{ti}!}{(N_{ti} - N_{ai})!\, N_{ai}!}$$

The probability that $N_{ai}$ or fewer accidents will be observed out of $N_{ti}$ Bernoulli trials can be computed as follows:

$$P(X \le N_{ai}, N_{ti}; P_i) = \sum_{i=0}^{N_{ai}} \frac{N_{ti}!}{(N_{ti} - i)!\, i!} P_i^i (1 - P_i)^{N_{ti} - i}$$

The probability that $N_{ai}$ or more accidents will be observed is expressed as

$$P(X \geq N_{ai}, N_{ti}; P_i) = 1 - P[X \leq (N_{ai} - 1)]$$
$$= 1 - \sum_{i=0}^{N_{ai-1}} \frac{N_{ti}!}{(N_{ti} - i)!\, i!} P_i^i (1 - P_i)^{N_{ti} - i}$$
$$= P(SF_i / N_{ai})$$

if $P(X \geq N_{ai}, N_{ti}; P_i) \leq P_{cr}$

Where $P_{cr}$ is some established threshold for making a classification decision, then the feature vector $\boldsymbol{X_i}[\boldsymbol{X_{a1}}, \boldsymbol{X_{a2}}, \ldots, \boldsymbol{X_{an}}]$ is classified as not belonging to a specific SPF $SF_i$. For the purposes of screening the inventory of existing intersections for potential projects, $P_{cr}$ should not exceed 1%. In the framework of performing diagnostics on the projects already selected for design, $P_{cr}$ value of up to 5% can be considered. In terms of accident analysis, it means that a roadway segment or junction that generated $\boldsymbol{X_{a1}}$ $[\boldsymbol{X_{a1}}, \boldsymbol{X_{a2}}, \ldots, \boldsymbol{X_{an}}]$ contains an element which triggers deviation from a random statistical process in the direction of reduced safety.

## APPLICATION OF DIRECT DIAGNOSTICS METHODOLOGY

To illustrate this concept further, consider the following example. If a die is rolled 10 times and a six is observed eight times, what is the probability that we are rolling a fair die? Considering that Bernoulli probability of observing a six on any given roll is .167, the probability of observing it eight times or more is approaching 0.

To illustrate the application of the concept of direct diagnostics to accident analysis, examine a case history of diagnosing and addressing a safety problem at an urban signalized intersection. The accident distribution profile for this location is presented in Figure 1. A total of 246 accidents were reported in the 5-year period and 97 of them were approach-turn accidents. The approach-turn accident is the most frequent accident type at this location. The critical question in the accident analysis of this intersection can be formulated as follows: Is it normal to experience 97 approach-turn accidents out of 246 total, or is there something present at the site which triggers increased frequency of these accidents? Direct diagnostics analysis can help answer this question. Based on 8 years of records presented in Table 1, approach-turn accidents represent 19% of the total accidents at urban signalized intersections, then $P_i = .19$. Compute the probability of observing 97 or more approach-turn accidents if 246 accidents have occurred:

$$P(X \leq x) = B(x, n; p) = \sum_{i=0}^{x} \frac{n!}{(n - i)!\, i!} p^i (1 - p)^{n-i}$$

The probability of observing 97 or more approach-turn accidents out of 246 total accidents at a "normal" urban signalized intersection is approaching 0, which suggests that there is a significant potential for accident reduction.

$$P(X \geq 97) = 1 - P(X \leq 96)$$

$$P(X \geq 97) = 1 - \sum_{i=0}^{96} \frac{246!}{(246 - i)!\, i!} (.19^i)(1 - .19)^{246-i} \approx 0$$

In other words, there is something in the environment of this intersection that triggers deviation from the random statistical process in the direction of reduced safety. Review of the collision

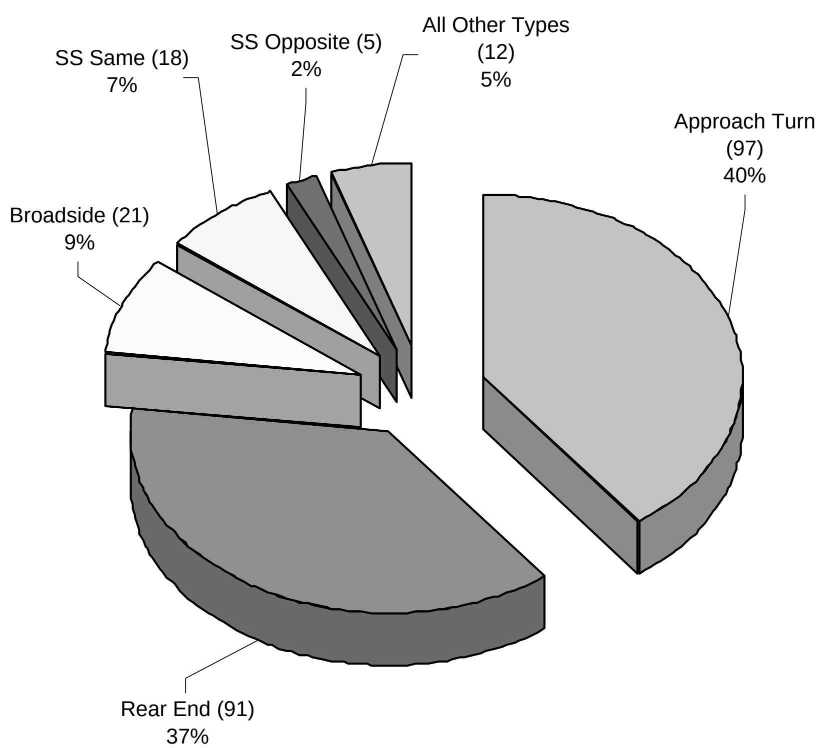


**FIGURE 1 Accident distribution profile at Site 1 (urban, signalized) (SS = sideswipe).**

**TABLE 1 Diagnostic Norms for Urban and Rural Intersections**

| Urban 4-Lane Divided Signalized Intersections | | | |
|---|---|---|---|
| **Description** | **Accidents** | **Percent** | |
| PDO | 32,885 | 67.79% | |
| INJ | 15,507 | 31.97% | |
| FAT | 117 | 0.24% | 100.00% |
| Persons Injured | 24,035 | N/A | |
| Persons Killed | 121 | N/A | |
| Single Vehicle Accidents | 3,389 | 6.99% | |
| Two Vehicle Accidents | 39,046 | 80.49% | |
| Three or more Vehicle Accidents | 5,928 | 12.22% | |
| Unknown Number of Vehicles | 146 | 0.30% | 100.00% |
| Overturning | 202 | 0.42% | |
| Other Non Collision | 187 | 0.39% | |
| School Age Pedestrians | 103 | 0.21% | |
| All Other Pedestrians | 832 | 1.72% | |
| Broadside | 7,742 | 15.96% | |
| Head On | 291 | 0.60% | |
| Rear End | 21,565 | 44.46% | |
| Sideswipe (Same Direction) | 3,678 | 7.58% | |
| Sideswipe (Opposite Direction) | 277 | 0.57% | |
| Approach Turn | 9,329 | 19.23% | |
| Overtaking Turn | 750 | 1.55% | |
| Parked Motor Vehicle | 768 | 1.58% | |
| Railway Vehicle | 3 | 0.01% | |
| Bicycle | 623 | 1.28% | |
| Motorized Bicycle | 5 | 0.01% | |
| Domestic Animal | 16 | 0.03% | |
| Wild Animal | 83 | 0.17% | |
| Total Fixed Objects | 1,895 | 3.91% | |
| Total Other Objects | 101 | 0.21% | 100.00% |
| Daylight | 35,223 | 72.61% | |
| Dawn or Dusk | 1,589 | 3.28% | |
| Dark - Lighted | 10,420 | 21.48% | |
| Dark - Unlighted | 703 | 1.45% | |
| Unknown Lighting | 574 | 1.18% | 100.00% |
| No Adverse Weather | 43,150 | 88.95% | |
| Rain | 2,502 | 5.16% | |
| Snow or Sleet or Hail | 2,006 | 4.14% | |
| Fog | 69 | 0.14% | |
| Dust | 5 | 0.01% | |
| Wind | 160 | 0.33% | |
| Unknown Weather | 617 | 1.27% | 100.00% |
| Dry Road | 40,916 | 84.35% | |
| Wet Road | 4,413 | 9.10% | |
| Muddy Road | 84 | 0.17% | |
| Snowy Road | 789 | 1.63% | |
| Icy Road | 985 | 2.03% | |
| Slushy Road | 185 | 0.38% | |
| Foreign Material Road | 119 | 0.25% | |
| With Road Treatment | 172 | 0.35% | |
| Dry with Icy Road Treatment | 11 | 0.02% | |
| Wet with Icy Road Treatment | 5 | 0.01% | |
| Snowy with Icy Road Treatment | 0 | 0.00% | |
| Icy with Icy Road Treatment | 4 | 0.01% | |
| Slushy with Icy Road Treatment | 1 | 0.00% | |
| Unknown Road Condition | 825 | 1.70% | 100.00% |
| **Total Accidents:** | **48,509** | | |
| **Total Number of Locations:** | **4,450** | | |

| Rural 2-Lane Undivided Unsignalized Intersections | | | |
|---|---|---|---|
| **Description** | **Accidents** | **Percent** | |
| PDO | 7,485 | 61.76% | |
| INJ | 4,460 | 36.80% | |
| FAT | 175 | 1.44% | 100.00% |
| Persons Injured | 7,426 | N/A | |
| Persons Killed | 210 | N/A | |
| Single Vehicle Accidents | 4,473 | 36.91% | |
| Two Vehicle Accidents | 7,051 | 58.18% | |
| Three or more Vehicle Accident | 567 | 4.68% | |
| Unknown Number of Vehicles | 29 | 0.24% | 100.00% |
| Overturning | 1,044 | 8.61% | |
| Other Non Collision | 235 | 1.94% | |
| School Age Pedestrians | 8 | 0.07% | |
| All Other Pedestrians | 59 | 0.49% | |
| Broadside | 2,098 | 17.31% | |
| Head On | 256 | 2.11% | |
| Rear End | 2,650 | 21.86% | |
| Sideswipe (Same Direction) | 336 | 2.77% | |
| Sideswipe (Opposite Direction) | 285 | 2.35% | |
| Approach Turn | 539 | 4.45% | |
| Overtaking Turn | 919 | 7.58% | |
| Parked Motor Vehicle | 333 | 2.75% | |
| Railway Vehicle | 9 | 0.07% | |
| Bicycle | 68 | 0.56% | |
| Motorized Bicycle | 0 | 0.00% | |
| Domestic Animal | 219 | 1.81% | |
| Wild Animal | 547 | 4.51% | |
| Total Fixed Objects | 2,374 | 19.59% | |
| Total Other Objects | 124 | 1.02% | 100.00% |
| Daylight | 8,234 | 67.94% | |
| Dawn or Dusk | 548 | 4.52% | |
| Dark - Lighted | 681 | 5.62% | |
| Dark - Unlighted | 2,478 | 20.45% | |
| Unknown Lighting | 179 | 1.48% | 100.00% |
| No Adverse Weather | 10,169 | 83.90% | |
| Rain | 439 | 3.62% | |
| Snow or Sleet or Hail | 1,058 | 8.73% | |
| Fog | 116 | 0.96% | |
| Dust | 6 | 0.05% | |
| Wind | 151 | 1.25% | |
| Unknown Weather | 181 | 1.49% | 100.00% |
| Dry Road | 9,265 | 76.44% | |
| Wet Road | 825 | 6.81% | |
| Muddy Road | 16 | 0.13% | |
| Snowy Road | 488 | 4.03% | |
| Icy Road | 988 | 8.15% | |
| Slushy Road | 166 | 1.37% | |
| Foreign Material Road | 34 | 0.28% | |
| With Road Treatment | 52 | 0.43% | |
| Dry with Icy Road Treatment | 3 | 0.02% | |
| Wet with Icy Road Treatment | 0 | 0.00% | |
| Snowy with Icy Road Treatment | 2 | 0.02% | |
| Icy with Icy Road Treatment | 5 | 0.04% | |
| Slushy with Icy Road Treatment | 0 | 0.00% | |
| Unknown Road Condition | 276 | 2.28% | 100.00% |
| **Total Accidents:** | **12,120** | | |
| **Total Number of Locations:** | **51,814** | | |

NOTE: PDO = property damage only, INJ = injury, FAT = fatality.

diagram followed by the field investigation revealed that a double left turn at each approach could be performed during the permitted turn phase. Permitted left turn on green with double left-turn lane assignment is generally associated with limited sight distance and consequently a high number of approach-turn accidents. This sort of safety problem at a signalized intersection can be effectively addressed by introducing protected left-turn phasing only.

Another signalized intersection in the urban area experienced 112 accidents over a 3 year period. The accident distribution profile for this location is presented in Figure 2. Out of 112 accidents reported in the 3-year period, 36 were broadsides. Compute the probability of observing 36 or more broadside accidents if 112 accidents have occurred, considering that Bernoulli probability of broadside collisions at an urban signalized intersection (Table 1) is 15.96%.

$$P(X \leq x) = B(x, n; p) = \sum_{i=0}^{x} \frac{n!}{(n-i)!\,i!} p^i (1-p)^{n-i}$$

$$P(X \geq 36) = 1 - P(X \leq 35)$$

$$P(X \geq 36) = 1 - \sum_{i=0}^{35} \frac{112!}{(112-i)!\,i!} (.1596^i)(1-.1596)^{112-i} \approx .00$$

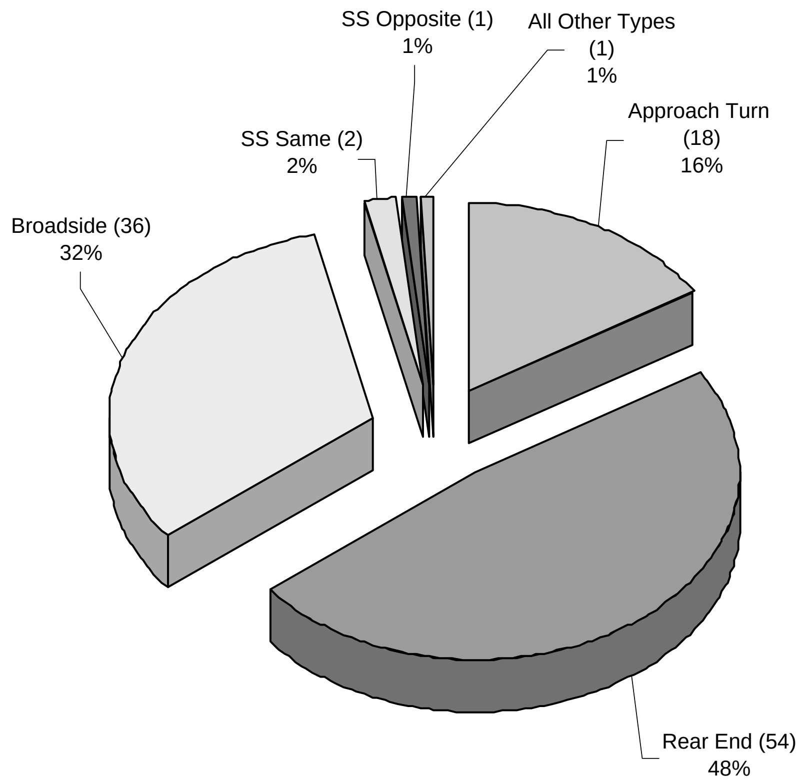


**FIGURE 2 Accident distribution profile at Site 2 (urban, signalized) (SS = sideswipe).**

The probability of observing 36 or more broadside accidents out of 112 total accidents within a random statistical process is extremely small, which suggests a high potential for accident reduction. Field investigation revealed that the visibility of a signal head on one of the approaches was obstructed by the foliage of an old tree. Trimming of the tree is expected to result in significant accident reduction at a minimal cost.

The accident type distribution profile at an unsignalized intersection in the rural area is presented on Figure 3. The probabilities of various accident types at rural unsignalized intersections and related parameters are included in Table 1. Although the number of accidents is not exceedingly high, the accident history reflects a well defined pattern of rear-end collisions. Considering that Bernoulli probability of rear-end collisions is 21.86%, the probability of observing 9 or more rear-end collisions out of 12 accidents at such a location is approaching 0.

$$P(X \geq 9) = 1 - P(X \leq 8)$$

$$P(X \geq 9) = 1 - \sum_{i=0}^{8} \frac{12!}{(12-i)!\,i!}(.2186^{i})(1-.2186)^{12-i} \approx 0$$

Review of a collision diagram followed by a field inspection revealed that a heavy left-turn movement at the approach where all of the rear-end collision occurred was not channelized. Construction of the left-turn bay at this approach is expected to reduce substantially the number of the rear-end collisions at this intersection.

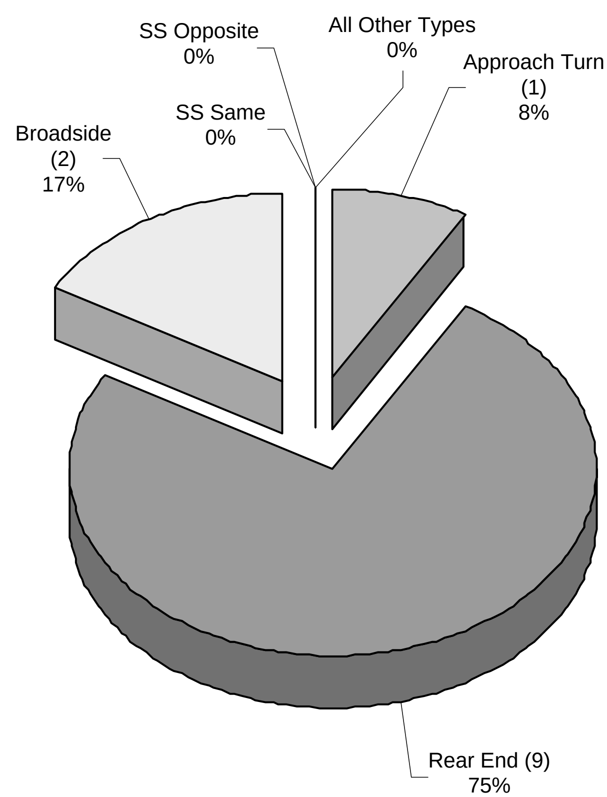


**FIGURE 3 Accident distribution profile at Site 3 (rural, unsignalized) (SS = sideswipe).**

## SUMMARY

Direct diagnostics methodology is primarily intended for projects already selected for design, although it can also be used for screening purposes. Detection of accident patterns at an intersection should always be followed up by the preparation of a collision diagram, field visit, and plans review. In many cases overall accident frequency is well within expected range, while accident patterns susceptible to correction are still present. Detection of an accident pattern at an intersection suggests a presence of an element or elements in the roadway environment that triggered a deviation from a random statistical process in the direction of reduced safety. Identification of such an element always provides a critical clue to accident causality. In many cases, the expected or normal proportion of accidents is counterintuitive, which further emphasizes the need for the creation of a framework of diagnostic norms for various types of intersections. The fact that cross-road ADT is not available in most states often makes intersection screening using SPF impractical. With this in mind, direct diagnostics methodology offers a useful tool to practitioners; at the same time diagnostic norms for accident analysis should be used in the same way as the *Diagnostic and Statistical Manual of Mental Disorders,* as guidelines to be informed by professional judgment and not in a cookbook fashion.

## REFERENCES

1. Hauer, E. Identification of Sites with Promise. In *Transportation Research Record 1542,* TRB, National Research Council, Washington, D.C., 1996, pp. 54–60.
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders,* 1952.
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders,* 4th ed., 1994.

*Publication of this paper sponsored by Committee on Safety Data, Analysis, and Evaluation.*

**TRANSPORTATION RESEARCH BOARD**
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

THE NATIONAL ACADEMIES™

*Advisers to the Nation on Science, Engineering, and Medicine*

The nation turns to the National Academies—National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council—for independent, objective advice on issues that affect people's lives worldwide.

www.national-academies.org