Development of a Road Safety Risk Index

Article *in* Transportation Research Record Journal of the Transportation Research Board · January 2002

DOI: 10.3141/1784-05

Tarek Sayed
University of British Columbia - Vancouver

Paper No. 02-2814

Development of a Road Safety Risk Index

Paul de Leur and Tarek Sayed

The cornerstone of most safety management programs consists of a "collision-prone location" program, where significant collision history must exist and be identified before improvements are recommended. Often, these programs are solely dependent on collision records and thus program success is governed by the data quality. Unfortunately, in many jurisdictions in North America, the quantity and quality of collision data have been degrading for several years. This growing problem is jeopardizing the success and continuance of many road safety programs. To help mitigate this problem, it is believed that a subjective evaluation technique could be developed that does not rely on collision statistics and that could be used to identify and diagnose problematic areas. The development and application of a risk index used for road safety evaluation are described. The risk index is developed as a driver-based, subjective assessment of the potential road safety risks for in-service roadways. The objective of developing the safety risk index is to produce a technique to support road safety analysis that does not rely on deteriorating collision data. The road safety risk index was developed and tested to ensure consistency between observers in their subjective assessment of safety. In addition, the results from the risk index were compared with results from objectively derived road safety measures to evaluate the success of the road safety risk index. The comparison indicates that there is a statistically significant agreement between the results of the risk index and the objectively derived road safety measures.

Given the importance of reducing the social and economic costs associated with collisions, most road authorities employ some type of road safety management program, designed to improve the road safety performance for the system users. The cornerstone of most safety management programs consists of a "collision-prone location" program. In these programs, a significant collision history must exist and be identified before road improvements are recommended.

Since a collision-prone location program is solely dependent on collision records, the success of the program is governed by the quality of the data. Unfortunately, in many jurisdictions, the quantity and quality of collision data have been degrading for several years. This problem can jeopardize the success and continuance of these programs.

To help mitigate these problems, it is believed that a subjective evaluation technique could be developed that does not rely on collision statistics and that could be used to identify and diagnose problematic areas. This paper describes the development and application of a risk index to be used for road safety evaluation. The risk index is developed as a driver-based, subjective assessment of the potential road safety risks for in-service roadways. The main objective of developing a road safety risk index is to produce a technique to support road safety analysis and a technique that does not rely on deteriorating collision data. Another reason for developing the subjective and driver-based evaluation of road safety risk is that there are important contributing factors that are useful but often cannot be extracted from a review of the collision statistics.

P. de Leur, Insurance Corporation of British Columbia, 305-2191 West 1st Avenue, Vancouver, BC V6K 1E7, Canada. T. Sayed, University of British Columbia, 2324 Main Mall, Vancouver, BC V6T 1Z4, Canada.

PREVIOUS WORK

In reviewing the literature, several sources describe different efforts to evaluate road safety subjectively. Two approaches are presented to contrast the different approaches and to demonstrate the benefits and limitations of each. One approach is the traffic conflict technique, and the second is a drive-through subjective rating system.

Traffic Conflict Technique

The concept of traffic conflicts was first proposed by Perkins and Harris (*1*) as an alternative to accident data. Their objectives were to define traffic incidents that occur frequently, can be clearly observed, and are related to accidents. A common definition of a traffic conflict is an observable situation in which two or more road users approach each other in space and time for such an extent that there is a risk of collision if their movements remain unchanged.

A variety of observation methods have been developed to measure traffic conflicts. These methods can be classified as subjective or objective. Subjective methods include considerable judgment by the conflict observer and are criticized by several researchers because the grading of severity of the evasive action can vary greatly from one observer to another. Objective methods include a cardinal or ordinal time-proximity dimension in the severity scale. The most widely used measure is the time to collision defined as the time for two vehicles to collide if they continue at their present speed and on the same path. However, it is noted that the traffic conflict technique requires the interaction of two vehicles and thus may not be useful for rural roads where many incidents are single-vehicle, off-road incidents.

Subjective Rating Systems

Another approach to subjective road safety evaluation involves a drive-through technique. A study completed by the Transport and Road Research Laboratory (*2*) investigated the impact of road design characteristics on driver perception and behavior and the propensity for driver risk acceptance. A 26-km route was selected for investigation and 60 drivers were used to make an assessment of the road safety risk at specific locations along the route.

The subjective safety rating was determined by having each test participant drive the route at a "comfortable," self-selected speed and then asked to give a rating of the road safety risk at 45 locations along the route. The rating was based on a subjective 11-point scale, with a score of 0 representing "no chance of a near miss" and a score of 10 representing a "good chance of a near miss." The objective safety rating was determined by calculating the accident rate based on historical collision statistics. In addition, a driver's selection of vehicle speed was also recorded as it could be used to reflect the perceived safety risk between locations. Higher speeds would reflect lower perceived risk, whereas lower speeds indicate higher risk.

The subjective and objective scores were then used to rank the road-user risk at each location and then to compare results. One test was to determine the agreement between observers in ranking the 45 locations, and the second test was to compare the rank of the subjective risk scores with the objective risk scores. The agreement between drivers (Test 1) was found to be significant. For Test 2, the Spearman rank correlation coefficient was used to determine the level of agreement between the subjective and objective risk scores. The coefficient was 0.37, indicating that the agreement was significantly less than perfect (1.0 represents perfect correlation). The reduction in vehicle speed selection and a driver's adaptation to potential hazard may be the principle reason for the lack of correlation between subjective and objective risk ratings.

OBJECTIVES FOR ROAD SAFETY RISK INDEX

Several objectives were specified in the development of the driver-based, subjective risk index including the following:

1. Construct the process with quantifiable results to prioritize locations of high-risk.
2. Ensure that the process is replicable and independent of the observer.
3. Ensure that the process is flexible and could be tailored to individual projects.
4. Ensure validity by comparing with objective safety performance measures.

DEVELOPMENT OF ROAD SAFETY RISK INDEX

Concept for Road Safety Risk Index

In the conceptual development of a risk index, it was necessary to consider the fundamental elements that can describe road safety in a quantifiable manner. Many road safety-engineering researchers (*3–5*) have isolated three elements that can be used to define safety risk. Although there is some ambiguity over the terminology, the three fundamental elements used to describe road safety risk include exposure, probability, and consequence as defined as follows:

$$\text{Risk} = \text{function of (exposure, probability, and consequence)}$$

where

exposure = measure to quantify the "exposure" of road users to potential roadway hazards,
probability = measure to quantify the chance of a vehicle being involved in a collision, and
consequence = measure to quantify the severity level resulting from potential collisions.

Methodology for Road Safety Risk Index

In formulating the methodology for a road safety risk index (RSRI), a three-step sequential process was established.

Step 1: Identify Factors to Be Considered for RSRI

When determining road features that impact safety and that should be used to develop the RSRI, it is often useful to separate the roadway into urban and rural environments. In most cases, each environment requires a focus on different road factors, as roadway design characteristics are considerably different between urban and rural environments (it is recognized that some road features may be common to both road environments). In addition, the ability to assess and collect risk information is different within each environment. In a rural environment, changes in road character are few, the driving task is generally simple, and thus the assessment and recording of risk are possible while driving at the posted speed. The opposite is true for an urban setting and risk assessment cannot be completed while driving.

Table 1 provides a list of roadway characteristics that may be considered in the assessment of road safety risk for each roadway environment. Not all factors listed in the table must be included in the formulation of the risk index and, alternatively, other characteristics not included may be added to the list.

Step 2: Formulate Guidelines for RSRI

Once the factors that impact road safety are selected, it is necessary to form guidelines to evaluate the road safety risk. For each factor, the elements of risk (i.e., exposure, probability, and consequence) are evaluated in formulating a risk score.

Exposure is evaluated by using the traffic volume that encounters a hazardous road feature. In a rural environment, generally the main line traffic volume is sufficient to quantify the exposure level. The exception may be at major intersections at which the traffic volume on the intersecting road may be important in determining the exposure to risk. In an urban environment, the traffic volume on both the major and minor roads is required, and in some cases the traffic volume by direction and/or lane movements is useful.

Probability is evaluated separately for each road feature to assess its potential in causing an incident. General guidance on how to assess the risk probability is provided for each factor listed in Table 2 and is based on the input and consensus from a group of road safety experts. However, it is important to recognize that each factor should be evaluated based on location-specific characteristics and, therefore, the guidance may deviate from what is provided herein.

In terms of the consequence element of risk, vehicle speed is one factor that influences the consequence of an incident and thus may be used to assess consequence. Alternatively, road features can provide surrogate measures for vehicle speed such as the higher speeds within passing zones and the road features causing high-speed differential. Thus the consequence level may also be evaluated based on various roadway features. In addition, the level of forgiveness of a roadside area is an important factor in determining the consequence in a rural road environment, as many collisions are single-vehicle, off-road crashes.

Step 3: Procedures to Obtain RSRI

In formulating the methodology for the RSRI several questions arose concerning how to quantify the components of risk during the drive-through review. The questions are presented as follows, together with a response for a solution for the road safety risk index.

1. *For each road feature, how should the component of risk be evaluated?* Based on the experience from the traffic conflict techniques

TABLE 1 Road Features Used to Formulate the Road Safety Risk Index

Environment	Road Feature	General Description of Risk
Rural	Horizontal and/or Vertical Curve	Locations where the horizontal or vertical alignment may cause safety concerns.
	Highway Access	High frequency of access points or access points that are particularly hazardous.
	Overtaking	Passing locations that are considered risky or unsafe, or a lack of opportunity to pass.
	Roadside Hazard	Locations where the roadside area creates a risk in the event of an off-road excursion.
	Road Surface/Super-elevation	Road surface conditions are problematic due to rutting, ponding, super-elevation.
	Design Consistency/Expectation	Inconsistent road features that may violate the driver's expectation.
Urban	Intersection Configuration	Intersection configuration or alignment may create potential safety problems.
	Traffic Control	Sites with no, or with inappropriate or poorly visible, traffic control devices.
	Roadway Access	Locations lacking access control or access points that impact on safety performance.
	Cross-sectional Elements	The cross-sectional elements of a road may give rise to road safety concerns.
	Road Friction/Maneuverability	Locations where road friction (i.e. parking) or poor maneuverability creates safety risk.
	Illumination and Road Markings	Illumination levels are low or road marking and signs are inappropriate or misleading.
	Road Surface Drainage	Sites where road surface condition is poor, evidenced by rutting, cracking, or ponding.

TABLE 2 Evaluating the Probability Component of Road Safety Risk

Environment	Road Feature	Evaluation of Probability
Rural	Horizontal and/or Vertical Curve	- radius/sharpness of curve(s) - presence of compound curves (s-curves) - combination of horizontal and vertical curves
	Highway Access	- access frequency/density - access alignment/connection - sight-distance from access location
	Overtaking	- length of passing zone - sight-distance in passing zone - opportunity for passing
	Roadside Hazard	- shoulder width and condition - degree of horizontal and vertical curve - conflict points from passing and/or access
	Road Surface & Super-elevation	- presence of rutting, ponding, cracking, holes - inappropriate super-elevation
	Design Consistency/Expectation	- unexpected feature requiring driver action - inconsistent road design features
Urban	Intersection Configuration	- oblique alignment of the intersection - level of channelization
	Traffic Control	- inappropriate and/or degree of traffic control - visibility of traffic control devices
	Roadway Access	- access frequency/density - access alignment/connection - sight-distance from access location
	Cross-sectional Elements	- narrow lane widths - facilities for alternate modes (sidewalks) - proximity of roadside hazards (i.e., poles)
	Road Friction/Maneuverability	- features creating road friction (i.e., parking) - ability to maneuver with ease (change lanes)
	Illumination and Road Markings	- low or inappropriate level of illumination - consistent and clear signs and markings
	Road Surface/Drainage	- presence of rutting, ponding, cracking, holes - opportunity for poor drainage

and suggestions by Wedley for analytical hierarchy of qualitative information (*6*), it was felt that four risk levels were appropriate. A score of 3 would represent high risk, a 2 would represent moderate risk, a 1 would represent minor risk, and a score of 0 could be used to indicate that the roadway feature poses no risk.

2. *Can each risk level be definitively distinguished during a drive-through?* There must be adequate opportunity to differentiate between the four risk levels. Guidelines associated with each factor are defined in a way that the road safety risk is assigned in a relative and subjective manner.

3. *Are some road features more important to safety than others?* At times, a specific road feature may be considered more hazardous than another (e.g., the poor horizontal curves are more problematic than the access points). If this is the case, a weighting factor can be applied to the risk scores for the problematic road feature to reflect this risk unbalance.

Formulation of RSRI

The exposure component of the RSRI can be determined in two ways. When a location is being studied in relation to a large reference population, specific volume levels can be developed to identify high, medium, and low exposure levels. Alternatively, when a specific corridor or group of locations is being investigated, exposure can be determined relative to the corridor or group of locations by using traffic volume data as shown in Equations 1 and 2 for urban and rural environments, respectively. This provides an exposure score ranging from zero to a maximum of 3.0, with a high score representing high exposure.

$$\text{Exposure}_{\text{Urban}} = \left(\frac{V_{i(\text{mjr})} \times V_{i(\text{mnr})}}{V_{\max(\text{mjr})} \times V_{\max(\text{mnr})}}\right) \times 3.0 \qquad (1)$$

$$\text{Exposure}_{\text{Rural}} = \left(\frac{V_i}{V_{\max}}\right) \times 3.0 \qquad (2)$$

where

$V_{\max(\text{mjr})}$ = maximum volume on the major road,
$V_{\max(\text{mnr})}$ = maximum volume on the minor road,
$V_{\max}$ = maximum volume on the corridor under review, and
V_i = volume at the location of a specific road feature.

The probability component of risk is obtained by using the guidelines provided in Table 2 and by making an assessment of each road feature using the four-point scale. This provides a probability score for each road feature ranging from 0 to 3.0, with a high score representing a high probability of an incident.

Several factors are used to gauge the consequence such as vehicle speed, potential for speed differential, mix of vehicle sizes, and roadside hazards. Thresholds are established for each of these factors and are used in combination to define consequence levels. The factors can be averaged to determine the composite consequence rating. Rather than identifying specific ranges for each factor, it is possible to calculate the relative consequence by using Equation 3 (using posted speed as the example), thus providing the results in the four-point scale (0 to 3).

$$\text{Consequence}_{\text{Rural}} = \left(\frac{PS_i}{PS_{\max}}\right) \times 3.0 \qquad (3)$$

where PS_i is the posted speed at the location of a specific feature and $PS_{\max}$ is the maximum posted speed. ($PS_{\max}$ can be the maximum posted speed on the corridor under review to assess relative consequence level on the corridor, or $PS_{\max}$ can be based on a larger area to gauge the relative consequence level beyond the corridor under review.)

The concept of the RSRI was presented previously in this paper and the reader is directed to the references for the theoretical background of the concept. RSRI can be formulated in two ways. The first index, $\text{RSRI}_{\text{specific}}$, defines the risk associated with each road feature, obtained by combining the scores for the three components of risk. $\text{RSRI}_{\text{specific}}$ identifies problem sites and facilitates the application of known improvements to address specific road deficiencies. The second risk index, $\text{RSRI}_{\text{combined}}$, defines overall risk by combing the $\text{RSRI}_{\text{specific}}$ scores for all road features. The formulations are shown in Equations 4 and 5.

$$\text{RSRI}_{\text{specific}} = E_i \times P_i \times C_i \qquad (4)$$

$$\text{RSRI}_{\text{combined}} = \sum_{i=1}^{n} E_i \times P_i \times C_i \qquad (5)$$

where

E_i = risk score due to exposure for road feature *i*,
P_i = risk score due to probability for road feature *i*,
C_i = risk score due to consequence for road feature *i*, and
n = number of road features investigated.

CASE STUDY APPLICATION OF RSRI

Location and Background Information

The corridor selected to test the RSRI was the Trans Canada Highway (TCH) in British Columbia between the city of Kamloops and the Alberta border. The corridor is 370 km in length and is the principle east-west highway connecting the city of Vancouver with the rest of Canada. The corridor is mainly rural, but passes through many small communities and was selected for several reasons. The character of the road is highly variable with areas of both generous and compromised road design standards due, in part, to the topographical constraints along the route. In addition, the British Columbia Ministry of Transportation had identified the route for a major upgrading and, as such, it was believed that the RSRI might be useful for the Ministry in assessing road safety.

Data Collection for RSRI for TCH Corridor

The corridor was first reviewed using the Ministry of Transportation's photo-log system, which provides an opportunity to view photographic images to understand the character of the corridor and to determine which road features will be investigated during the drive-through review. A meeting with Ministry staff was also conducted to gain an understanding of the hazardous features along the corridor.

After identifying the road features of concern on the corridor, it was determined that three factors would be investigated to test the RSRI. These factors would adequately demonstrate the process, which could then be applied in a similar manner to other factors and other corridors. The drive-through review was conducted in October 1998, when the conditions were favorable to observe the road features and a three-person team was used to assess and record the road safety risk, which was completed within 4 days. The three road features targeted for the case study were

1. Problems associated with access points onto the highway,
2. Problems with the opportunity/ability to complete passing maneuvers, and
3. Problems associated with roadside hazards.

Development of RSRI for TCH Corridor

Risk Assignment for Access

The exposure component of risk was calculated throughout the corridor by using the main-line traffic volume and Equation 2. The three evaluation guidelines suggested in Table 2 were used to assess the probability component of risk for each access point, providing sufficient guidance such that an observer could make a subjective judgment and assign a road safety risk score. In areas in which there were many access points, it was necessary to slow down or stop, in order that each access point could be adequately evaluated. The potential risk associated with each access point was evaluated by subjectively assessing the following:

- The frequency and/or density of access locations on the highway corridor,
- The horizontal and vertical alignment of the access road's connection to the highway, and
- The turning sight distance that is available to motorists who are attempting to gain access to the highway.

The posted speed and two other road features were used to assess the consequence component of risk for roadway access. The other factors included whether the access point is located in a passing section or in an area in which the roadside safety risk was high. These risk scores were being collected concurrently by the other observers. It is noted that the assessment of consequence risk was calculated in a similar manner for both the passing opportunity and roadside hazard.

Risk Assignment for Passing Opportunities

The exposure component of risk associated with each passing opportunity was calculated on the corridor by using the main-line traffic volume and Equation 2. Each passing opportunity on the test corridor was reviewed with respect to the probability of causing an incident. This included all passing zones, passing lanes, and climbing lanes. In assessing the safety risk for passing opportunity, three guidelines (as suggested in Table 2) were used, which provided sufficient guidance to the observer to assess subjectively the safety risk using the four-point ranking system. These guidelines include

- The overall length of the passing zone,
- The sight distance that is available while motorists attempt to complete a passing maneuver, and
- The frequency and duration of passing opportunities available along the corridor.

Risk Assignment for Roadside Hazard

Again, the exposure component of risk associated with the roadside hazard was calculated by using the main-line traffic volume and Equation 2. The three evaluation guidelines proposed in Table 2 were used to determine the probability component of risk for the roadside. These three factors provided adequate guidance for the observer to make a subjective assessment using the four-point scale. These factors include

- The degree of horizontal and vertical curves that may cause a vehicle to leave the roadway and enter the roadside area;
- The shoulder width, the surface type of the shoulder, and shoulder condition; and
- The frequency of conflict points that may create the potential for an off-road excursion.

Results of RSRI for TCH Corridor

The road safety risk index was calculated based on Equation 4 for each specific road feature and Equation 5 for the combined index. The results are summarized based on a total of 74 segments that are 5.0 km in length and are considered homogeneous based on roadway design and traffic characteristics. The corridor segmentation was completed with the assistance and knowledge from local staff from the Ministry of Transportation. The results of the risk index are shown in Table 3.

SUCCESS AND VALIDITY OF RSRI

Comparing RSRI Scores between Observers

An important success factor of the RSRI is the reliability of the risk assessment and data collection process. The reliability of the process can be measured by the similarity of the results produced by different observers for the same section of roadway. The inability to record similar safety risk produced by two different users will jeopardize the validity of the process. To address this concern, the level of agreement between observers was determined by using an 85-km segment from the study corridor. A second observer would make an assessment of the risk for the three factors (access, passing, and roadside), given the same guiding information provided in Table 2.

To evaluate the reliability of the risk index for access, 77 access points on the segment were used in the evaluation. To evaluate the reliability of the risk assignment for passing, 72 passing zones in both directions were included. Finally, to evaluate the reliability of risk assignment for roadside hazards, the segment was divided into 85 1-km segments with a risk score produced for each 1-km section.

To determine the consistency and therefore the reliability of the risk assignment between observers, the kappa statistic was used (*7, 8*). The kappa statistic (κ) provides a measure of the agreement between the two observers. The kappa statistic utilizes the subjective risk assignment score produced by each observer, categorized into three categories: high risk (score 3), moderate risk (score 2), and minor risk (score 1). The kappa statistic is defined as follows:

$$\kappa = \left(\frac{P - P_e}{1 - P_e}\right) \tag{6}$$

TABLE 3 Results of RSRI for TCH Corridor

Segment Number	Segment Length (Km)	Segment Volume (vpd)	Risk Index Score			
			Access	Passing	Roadside	Total
1	5.0	21608	1.05	10.01	11.00	22.06
2	5.0	14124	0.14	4.41	5.25	9.80
3	5.0	6832	0.11	5.30	7.06	12.47
4	5.0	6039	1.40	10.28	34.95	46.63
5	5.0	6039	0.40	0.96	1.28	2.65
6	5.0	6039	1.22	4.58	18.49	24.30
7	5.0	6039	2.66	6.18	15.07	23.92
8	5.0	6357	0.81	4.55	10.59	15.96
9	5.0	6639	2.05	4.89	11.69	18.63
10	5.0	5922	0.45	1.12	4.53	6.10
11	5.0	8986	2.74	1.93	29.77	34.44
12	5.0	10941	8.14	11.39	37.28	56.81
13	5.0	10941	6.32	4.72	7.17	18.21
14	5.0	10941	15.87	10.03	35.83	61.73
15	5.0	10941	6.66	3.40	16.65	26.72
16	5.0	10982	4.88	6.26	9.88	21.02
17	5.0	10982	0.52	0.00	5.23	5.76
18	5.0	10660	0.14	0.66	2.81	3.61
19	5.0	4311	0.68	1.11	5.72	7.51
20	5.0	4311	2.07	1.01	7.02	10.10
21	5.0	4379	1.09	0.78	4.58	6.44
22	5.0	5780	0.14	0.09	0.21	0.44
23	5.0	5546	0.56	2.19	2.75	5.50
24	5.0	5426	1.50	1.30	2.45	5.25
25	5.0	5291	0.15	0.15	0.00	0.30
26	5.0	5102	0.23	0.35	0.45	1.04
27	5.0	5021	0.28	0.42	0.61	1.31
28	5.0	4874	2.30	2.01	4.96	9.27
29	5.0	4874	0.25	0.38	3.45	4.08
30	5.0	4874	0.13	0.24	1.55	1.93
31	5.0	4804	1.03	1.19	2.02	4.24
32	5.0	4711	1.51	0.99	5.92	8.43
33	5.0	5688	1.00	3.34	6.11	10.45
34	5.0	6076	1.12	1.84	4.18	7.14
35	5.0	4976	0.28	0.31	0.81	1.40
36	5.0	4636	0.52	2.08	6.78	9.37
37	5.0	4646	0.57	3.40	7.57	11.54
38	5.0	4646	0.23	0.34	0.90	1.48
39	5.0	4646	0.25	0.37	1.73	2.34
40	5.0	4707	0.30	0.08	3.17	3.55
41	5.0	4718	0.05	1.24	5.34	6.63
42	5.0	4718	0.03	1.17	3.67	4.88
43	5.0	4742	1.24	0.23	6.16	7.63
44	5.0	4791	0.13	0.74	2.29	3.15
45	5.0	4791	0.15	1.24	2.71	4.11
46	5.0	4791	0.11	0.45	1.74	2.30
47	5.0	4791	0.10	0.79	1.10	1.99
48	5.0	4791	0.33	0.66	5.94	6.92
49	5.0	4791	0.82	2.50	9.67	12.99
50	5.0	4791	0.04	0.73	1.53	2.29
51	5.0	4791	0.12	0.69	1.95	2.77
52	5.0	4791	0.45	2.41	3.48	6.34
53	5.0	4791	0.36	2.42	5.68	8.46
54	5.0	4791	0.53	5.77	6.36	12.66
55	5.0	4791	0.26	2.28	4.83	7.37
56	5.0	4791	0.37	2.05	14.50	16.92
57	5.0	4791	0.76	3.99	18.66	23.41
58	5.0	4859	1.20	4.82	9.64	15.65
59	5.0	4863	1.72	5.39	12.44	19.55
60	5.0	4863	2.60	3.23	6.37	12.21
61	5.0	4863	2.12	1.95	15.06	19.13
62	5.0	4784	0.38	3.71	16.01	20.11
63	5.0	4296	1.78	0.00	4.18	5.96
64	5.0	4052	1.40	0.58	33.06	35.05
65	5.0	4052	0.91	3.28	11.48	15.66
66	5.0	4052	0.91	2.41	4.38	7.69
67	5.0	3950	0.51	1.19	1.07	2.77
68	5.0	3899	0.01	0.04	0.21	0.26
69	5.0	3899	0.01	0.16	0.55	0.73
70	5.0	3899	0.01	0.44	0.76	1.21
71	5.0	3899	0.18	0.24	1.66	2.08
72	5.0	3899	0.13	0.55	1.04	1.72
73	5.0	3899	0.05	0.05	0.00	0.09
74	5.0	3899	0.02	0.02	0.00	0.04

where P is overall percent agreement (0.0–1.0) and P_e is overall percent agreement expected by chance (0.0–1.0).

The overall percent agreement (P) is calculated by summing the number of observations in which agreement exists divided by the total number of observations. The overall percent agreement that can be expected by chance (P_e) can be calculated by the percentage of assignment by each observer into the number of categories. A positive value for the kappa statistic indicates agreement between observers, a value of zero represents the level of agreement that could be expected by chance, and a negative kappa indicates disagreement between observers. The variance of the kappa statistic is required to determine agreement between observers. The variance of kappa is calculated as follows:

$$\mathrm{Var}(\kappa) = \frac{1}{N} \times \frac{\left[\sum p_j^2 - \left(\sum p_j^2\right)^2\right]}{\left(1 - \sum p_j^2\right)^2} \tag{7}$$

where

N = total number of observations,
J = number of categories of classification, and
p_j = proportion of all assignments of the jth category.

Under a hypothesis of no agreement beyond chance, a value can be calculated to test the significance level of the agreement between the two observers. This value, defined as $k/\sqrt{\mathrm{Var}(\kappa)}$, is approximately distributed as a standard normal variant (*8*) and can be compared to the critical z-value to determine the level of significance.

To demonstrate, consider the road safety risk scores of 3, 2, or 1 assigned to the 77 access locations from the two observers as shown in Table 4. The overall percent agreement (P) is calculated by summing the values along the diagonal divided by the total number of observations. The overall percent agreement between observers that can be expected by chance (P_e) can also be calculated, as well as the kappa statistic, as follows:

$$P = \frac{9 + 14 + 26}{77} = 0.636$$

$$P_e = \left(\frac{14}{77} \times \frac{14}{77}\right) + \left(\frac{24}{77} \times \frac{29}{77}\right) + \left(\frac{39}{77} \times \frac{34}{77}\right) = 0.374$$

$$\kappa = \frac{0.636 - 0.374}{1 - 0.374} = 0.419$$

The variance of the kappa statistic is 0.0077 as calculated from Equation 7. To determine the level of agreement between observers and the level of significance, the ratio of $k/\sqrt{\mathrm{Var}(\kappa)}$ is calculated as 4.76. This value is then compared to the critical z-value of 2.33, representing the 99% significance level. Since 4.78 > 2.33, it can be concluded that agreement between observers is achieved and the process to assign road safety risk for access locations is valid. A similar reliability analysis was completed for the risk scores for both passing and roadside hazard. The subjective measures evaluated included the passing sight distance, the shoulder width, the degree of embankment slope, and the hazard level for objects in the roadside. The results are shown in Table 5.

The results indicate that the level of agreement between observers for three of the four subjective measures exceeded the 99% level of significance. The exception was that of the roadside shoulder where the significance test failed. The reason for the failure has less to do with the level of agreement between observers, which was very high ($P = 0.894$), but more as a result of the percent expected by chance ($P_e = 0.814$), which was similarly high. The reason for the high value of P_e is that the shoulder width and condition (the basis of risk assessment) do not vary significantly on the corridor.

Comparing Ranks with Objective Measures (Collision History)

In order to evaluate the success and validity of the RSRI, a meaningful road safety performance indicator based on collision data was used. The indicator used is defined as the potential for improvement (PFI), measured as the difference between the existing collision frequency and the expected collision frequency at a location. The expected collision frequency is determined by applying a valid collision prediction model and refined by applying the empirical Bayes (EB) technique (*9*). The difference between the existing and expected collision frequencies will facilitate the ranking of sites (a greater difference having a higher rank) and this rank will be compared with the ranking determined by the RSRI.

The starting point for this evaluation was the development of a collision prediction model to determine the expected collision frequency at each site. The technique used to develop the collision prediction model is based on the generalized linear modeling approach (GLIM). Several researchers [e.g., Hauer et al. (*10*) and Miaou and Lum (*11*)] have demonstrated the inappropriateness of conventional linear regression for modeling discrete, nonnegative, and rare events such as traffic collisions. GLIM has the advantage of overcoming these shortcomings associated with conventional linear regression.

Collision and traffic data were obtained for road segments of similar character to the roadway used in the RSRI case study. A model was then developed that predicts the 3-year collision frequency at a site based on the segment length, measured in kilometers (L), and the main-line traffic volume given by the average annual daily traffic (AADT) volume. The formulation of the model, the model parameters, and the indicators for the model significance, including the t-ratio, the k-value, and the Pearson χ^2 is provided

TABLE 4 Level of Agreement Between Two Observers for Access

Observer Two Risk Score	Observer One Risk Score			Total
	'3'	'2'	'1'	
'3'	9	3	2	14
'2'	4	14	11	29
'1'	1	7	26	34
Total	14	24	39	77

TABLE 5 Level of Agreement for Other Subjective Measures

Calculated Value	Passing Sight Distance	Roadside: Shoulder Width/Condition	Roadside: Side-Slope Hazard	Roadside: Hazardous Objects
P	0.611	0.894	0.647	0.694
P_e	0.470	0.814	0.362	0.448
(κ)	0.267	0.430	0.447	0.446
(P_3)	0.0347	0.0176	0.0706	0.0882
(P_2)	0.0903	0.0059	0.1059	0.0882
(P_1)	0.2292	0.0765	0.4056	0.1941
(P_0)	0.6458	0.9000	0.4294	0.6294
$(P_3)^2$	0.00121	0.00031	0.00498	0.00779
$(P_2)^2$	0.00815	0.00003	0.01121	0.00779
$(P_1)^2$	0.05252	0.00585	0.16474	0.03768
$(P_0)^2$	0.41710	0.81000	0.18439	0.39616
$\Sigma(P_i)^2$	0.47897	0.81619	0.36532	0.44941
Var (κ)	0.0128	0.0527	0.0068	0.0096
$k/\sqrt{Var(\kappa)}$	2.36	1.89	5.43	4.55
Significance at 99% Level	**Yes 2.36 > 2.32**	**No 1.89 < 2.32**	**Yes 5.43 > 2.32**	**Yes 4.55 > 2.32**

below. The Pearson χ^2 and the *t*-ratios indicate that the model has a reasonable goodness of fit.

$$\text{Collisions / 3 years} = 0.001302 \times (L \times \text{AADT})^{0.9645}$$
$$t_0 = 5.1, t_1 = 3.4, k = 1.44, \text{Pearson } \chi^2 = 171,$$
$$\text{degrees of freedom} = 156, \chi^2(156, 0.05) = 179 \qquad (8)$$

where *k* is the negative binomial parameter.

With the developed collision prediction model, the next step was to calculate the expected collision frequency at sites along the corridor using Equation 8. A total of 74 sites were selected on the 370-km corridor, with each site having a fixed length of 5.0 km. These segments are used to present the results of the RSRI.

Collision data were extracted from the provincial accident database to determine the existing collision frequency at each of the 74 sites. It is noted that collision underreporting is not a problem on this corridor and the stability and quality of the collision data are considered good. These collision counts were then subjected to an EB refinement technique to obtain a better estimate of the existing safety performance, produced as follows:

$$EB_{\text{safety estimate}} = \left[\frac{E(\Lambda)}{\kappa + E(\Lambda)}\right] \times (k + \text{count}) \qquad (9)$$

where $E(\Lambda)$ is the predicted collision frequency (collisions / 3 years) and count is the observed collision frequency (collisions / 3 years).

With the existing collision frequency (subjected to the EB refinement) and the expected collision frequency at each location, it is then possible to determine the PFI for the 74 sites. The PFI is simply measured as the difference between the existing collision frequency and the expected collision frequency. The rank is established based on descending order of the difference between existing and expected collision frequency. This rank is then compared to the ranking of the risk scores from the RSRI, as presented in Table 3.

The Spearman rank correlation coefficient is used to determine the level of agreement between the RSRI and the PFI. The Spearman rank correlation coefficient is often used as a nonparametric alternative to a traditional coefficient of correlation and can be applied under general conditions (*12*). To calculate the Spearman rank correlation coefficient, it is necessary to segment the data sets and then rank the paired data sets in ascending or descending order.

A summary of the total RSRI scores, the existing collision frequency, and the expected collision frequency for each of the 74 sites is shown in Table 6. Also in the table is the ranking established for each of the 74 sites based on both the RSRI and the PFI. The paired data set is then used in calculating the Spearman rank correlation coefficient (ρ_s) as shown in Equation 10. A score of 1.0 represents perfect correlation and a score of zero indicates no correlation. An advantage of using ρ_s is that when testing for correlation between two sets of data, it is not necessary to make assumptions about the nature of the populations sampled.

$$\rho_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)} \qquad (10)$$

where *d* is the difference between ranks and *n* is the number of paired sets.

Under a null hypothesis of no correlation, the ordered data pairs are randomly matched and thus the sampling distribution of ρ_s has a mean of zero and the standard deviation (σ_s) as given in Equation 11.

$$\sigma_s = \frac{1}{\sqrt{(n-1)}} \qquad (11)$$

Since this sampling distribution can be approximated with a normal distribution even for relatively small values of *n*, it is possible to test the null hypothesis on the statistic given in Equation 11. This value can be compared to a critical *z*-value of 2.33 representing the 99% level of significance.

$$z = \rho_s\sqrt{(n-1)} \qquad (12)$$

The results from the correlation analysis indicate that the ranking from the subjective RSRI and the objective PFI do agree at the 99% level of significance. With *n* equal to 74, the Spearman rank

TABLE 6 Summary of RSRI Score and Collisions on Segmented Corridor

Segment Number	Segment Length (Km)	RSRI Score	RSRI Rank	Collision Frequency (collisions / 3 yrs) Count	EB Estimate	Expected	PFI EB Estimate – Expected	PFI Rank	Rank Difference
1	5.0	22.06	10	197	195.5	93.20	102.32	2	8
2	5.0	9.80	28	83	82.6	61.85	20.70	45	-17
3	5.0	12.47	23	35	34.8	30.70	4.12	71	-48
4	5.0	46.63	3	51	49.9	27.25	22.63	42	-39
5	5.0	2.65	56	54	52.7	27.25	25.49	38	18
6	5.0	24.30	7	70	68.0	27.25	40.74	16	-9
7	5.0	23.92	8	65	63.2	27.25	35.98	21	-13
8	5.0	15.96	18	50	49.0	28.64	20.41	46	-28
9	5.0	18.63	15	72	70.2	29.86	40.33	17	-2
10	5.0	6.10	42	24	24.1	26.75	-2.61	73	-31
11	5.0	34.44	5	58	57.4	39.98	17.43	54	-49
12	5.0	56.81	2	78	77.2	48.35	28.85	30	-28
13	5.0	18.21	16	63	62.6	48.35	14.26	58	-42
14	5.0	61.73	1	109	107.4	48.35	59.02	9	-8
15	5.0	26.72	6	304	297.1	48.35	248.76	1	5
16	5.0	21.02	11	151	148.2	48.52	99.72	3	8
17	5.0	5.76	44	66	65.5	48.52	17.01	55	-11
18	5.0	3.61	51	34	34.4	47.15	-12.79	74	-23
19	5.0	7.51	35	64	61.2	19.69	41.49	15	20
20	5.0	10.10	27	60	57.4	19.69	37.74	20	7
21	5.0	6.44	40	54	51.9	19.99	31.87	25	15
22	5.0	0.44	70	26	26.0	26.13	-0.12	72	-2
23	5.0	5.50	45	45	44.0	25.11	18.89	50	-5
24	5.0	5.25	46	59	57.2	24.58	32.64	24	22
25	5.0	0.30	71	45	43.9	23.99	19.90	49	22
26	5.0	1.04	68	35	34.4	23.16	11.19	63	5
27	5.0	1.31	66	50	48.5	22.81	25.68	37	29
28	5.0	9.27	30	49	47.5	22.17	25.31	39	-9
29	5.0	4.08	50	78	74.8	22.17	52.65	11	39
30	5.0	1.93	62	64	61.6	22.17	39.45	18	44
31	5.0	4.24	48	46	44.6	21.86	22.75	41	7
32	5.0	8.43	32	57	54.9	21.45	33.46	23	9
33	5.0	10.45	26	90	86.8	25.72	61.09	8	18
34	5.0	7.14	37	34	33.7	27.42	6.28	68	-31
35	5.0	1.40	65	45	43.7	22.61	21.13	44	21
36	5.0	9.37	29	52	50.2	21.12	29.04	28	1
37	5.0	11.54	25	59	56.7	21.16	35.58	22	3
38	5.0	1.48	64	41	39.8	21.16	18.65	51	13
39	5.0	2.34	57	63	60.5	21.16	39.34	19	38
40	5.0	3.55	52	32	31.4	21.43	9.95	65	-13
41	5.0	6.63	39	70	67.2	21.48	45.67	14	25
42	5.0	4.88	47	52	50.2	21.48	28.73	31	16
43	5.0	7.63	34	50	48.3	21.59	26.75	35	-1
44	5.0	3.15	53	55	53.1	21.80	31.28	26	27
45	5.0	4.11	49	84	80.4	21.80	58.60	10	39
46	5.0	2.30	58	72	69.1	21.80	47.29	13	45
47	5.0	1.99	61	51	49.3	21.80	27.51	33	28
48	5.0	6.92	38	43	41.8	21.80	19.97	47	-9
49	5.0	12.99	21	28	27.6	21.80	5.84	69	-48
50	5.0	2.29	59	43	41.8	21.80	19.97	48	11
51	5.0	2.77	55	41	39.9	21.80	18.09	53	2
52	5.0	6.34	41	48	46.5	21.80	24.68	40	1
53	5.0	8.46	31	29	28.6	21.80	6.78	66	-35
54	5.0	12.66	22	72	69.1	21.80	47.29	12	10
55	5.0	7.37	36	29	28.6	21.80	6.78	67	-31
56	5.0	16.92	17	39	38.0	21.80	16.20	56	-39
57	5.0	23.41	9	90	86.1	21.80	64.25	7	2
58	5.0	15.65	20	36	35.2	22.10	13.11	59	-39
59	5.0	19.55	13	55	53.1	22.12	31.00	27	-14
60	5.0	12.21	24	35	34.3	22.12	12.15	62	-38
61	5.0	19.13	14	36	35.2	22.12	13.09	60	-46
62	5.0	20.11	12	106	101.1	21.77	79.35	6	6
63	5.0	5.96	43	111	105.2	19.63	85.53	5	38
64	5.0	35.05	4	125	117.8	18.55	99.28	4	0
65	5.0	15.66	19	49	46.9	18.55	28.40	32	-13
66	5.0	7.69	33	38	36.7	18.55	18.14	52	-19
67	5.0	2.77	54	42	40.4	18.10	22.25	43	11
68	5.0	0.26	72	24	23.6	17.87	5.70	70	2
69	5.0	0.73	69	29	28.2	17.87	10.35	64	5
70	5.0	1.21	67	35	33.8	17.87	15.93	57	10
71	5.0	2.08	60	31	30.1	17.87	12.21	61	-1
72	5.0	1.72	63	46	44.0	17.87	26.17	36	27
73	5.0	0.09	73	47	45.0	17.87	27.10	34	39
74	5.0	0.04	74	49	46.8	17.87	28.96	29	45

Notes: PFI = the Potential for Improvement is defined as the difference between the EB estimate (observed collisions) and the expected or normal collisions (from the prediction model). Difference in ranks between RSRI and potential for improvement.

correlation coefficient (ρ_s) was calculated to be 0.318 and the standard deviation (σ_s) was calculated to be 0.117. The calculated z-value was 2.66, thus indicating agreement at a 99% significance level when compared to a critical z-value of 2.33. These results provide further validation for the RSRI.

CONCLUSIONS

The RSRI process utilizes concepts related to the traffic conflict observation technique and other drive-through safety reviews. The reliability of the process stems from the well-defined and quantifiable characteristics of road features that are studied and scored while completing a drive-through review. These scores are combined to produce an overall road safety risk, formulated by combining three components of risk, namely, the exposure of road users to road hazards, the probability of becoming involved in a collision, and the resulting consequences should a collision occur.

A systematic process was described to determine which road features should be investigated and how each feature should be evaluated during the drive-through review. It was not reasonable or necessary to list all possible road features associated with the many different road types. Rather, several typical road features associated with rural and urban roads were provided to illustrate the process. This accommodated one objective: that the RSRI process is flexible and can adapt to the needs of many users and differing conditions.

By providing definitive guidelines on how to assess the various factors that influence road safety risk, the process is made replicable, with consistent results produced independent of the observer. The kappa statistic was used in the case study to determine the consistency of the road safety risk assignment between observers. The level of agreement was considered acceptable thereby supporting the requirement for a replicable RSRI process.

The validity of the RSRI was also evaluated by comparing the results of the risk index with an objective safety measure defined as the potential for improvement, based on the difference in the existing and expected collision frequencies. Accurate estimates of the existing and expected collision frequencies were obtained by using a collision prediction model and applying the empirical Bayes technique. The Spearman rank correlation coefficient was used to determine the agreement level between the results of the RSRI and the PFI. Sites were ranked according to both the RSRI and PFI, with the results of the Spearman correlation indicating agreement at a 99% significance level.

Due to the validation and quantifiable nature of the RSRI, the results can be used to support road safety analysis and decision making. In particular, intersections or road segments of high risk can be identified and isolated for road safety improvements. Another opportunity arising from the results of the RSRI is that specific road improvements can be formulated to address specific problems determined when each road feature is investigated separately. The drive-through process allows for the investigation of factors that are not normally available through a police collision report form, such as the potential for driver confusion. One final benefit associated with the production of the RSRI is that the drive-through process can assist in the formulation of appropriate road improvement strategies.

One of the fundamental limitations of the road safety risk index is the subjective nature of the process. Any process that relies on a subjective assessment can be susceptible to accuracy problems. Effort has been made to establish specific thresholds for the elements of the RSRI, attempting to make the process quantifiable, but it remains a subjective process and thus the accuracy can be questioned. In addition, it is recognized that, like many other road safety evaluation techniques, the RSRI may not adequately assess the human factor component inherent in road safety assessment.

REFERENCES

1. Perkins, S. R., and J. L. Harris. *Criteria for Traffic Conflict Characteristics.* Report GMR 632. General Motors Corporation, Warren, Mich., 1967.
2. *Aspects of Road Layout that Affect Drivers' Perception and Risk Taking.* TRR 602. Transport and Road Research Laboratory, 1990, pp. 122–127.
3. Hadden, W. Advances in the Epidemiology of Injuries as a Basis for Public Policy. *Public Health Reports,* Vol. 95, No. 5, 1980, pp. 411–421.
4. Hauer, E. Traffic Conflicts and Exposure. *Accident Analysis & Prevention,* Vol. 14, 1982, pp. 352–362.
5. Koornstra, M. J. The Evolution of Road Safety and Mobility. *IATSS Research,* Vol. 16, No. 2, 1992, pp. 129–148.
6. Wedley, W. C. Combining Qualitative and Quantitative Factors: An Analytic Hierarchy Approach. *Socio-Economic Planning Science,* Vol. 24, No. 1., 1990, pp. 57–64.
7. Cohen, J. A. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement,* Vol. 20, 1960, pp. 37–46.
8. Fleiss, J. L. Measuring the Effective Nominal Scale Agreement from Among Raters. *Psychological Bulletin,* Vol. 75, No. 5, 1971, pp. 378–382.
9. Sayed, T., L. F. Rodríguez, and S. Feng. Accident Prediction Models for Signalized Intersections in British Columbia. Presented at the 1998 ITE Annual Meeting, Toronto, Aug. 9–13, 1998.
10. Hauer, E., J. C. N. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185,* TRB, National Research Council, Washington D.C., 1988, pp. 48–61.
11. Miaou, S., and H. Lum. Modeling Vehicle Accident and Highway Geometric Design Relationships. *Accident Analysis & Prevention,* Vol. 25, No. 6, 1993, pp. 689–709.
12. Freund, J. E. *Modern Elementary Statistics,* 6th ed. Prentice-Hall, Englewood Cliffs, N.J., 1994.

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