Developing a Methodology for Implementing Safety Improvements on Low-Volume Roads in Montana

Task V Report: Assessment of Existing Methodologies

Prepared by

Ahmed Al-Kaisy, PhD, PE – Professor & Program Manager

Kazi Tahsin Huda – Graduate Research Assistant

of the

Western Transportation Institute
College of Engineering
Montana State University
Bozeman, MT 59717-4250

A report prepared for the

MONTANA DEPARTMENT OF TRANSPORTATION

in cooperation with the

U.S. DEPARTMENT OF TRANSPORTATION

FEDERAL HIGHWAY ADMINISTRATION

September 2020

DISCLAIMER STATEMENT

This document is disseminated under the sponsorship of the Montana Department of Transportation (MDT) and the United States Department of Transportation (USDOT) in the interest of information exchange. The State of Montana and the United States assume no liability for the use or misuse of its contents.

The contents of this document reflect the views of the authors, who are solely responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the views or official policies of MDT or the USDOT.

The State of Montana and the United States do not endorse products of manufacturers.

This document does not constitute a standard, specification, policy or regulation.

ALTERNATIVE FORMAT STATEMENT

MDT attempts to provide accommodations for any known disability that may interfere with a person participating in any service, program, or activity of the Department. Alternative accessible formats of this information will be provided upon request. For further information, call 406/444.7693, TTY 800/335.7592, or Montana Relay at 711.

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

This interim report summarizes the results of Task 5 for the project titled "Developing a Methodology for Implementing Safety Improvements on Low-Volume Roads in Montana." This task is intended to analyze and assess the various network screening approaches that are reported in the existing literature or included in the results of the practice survey conducted in the previous task. The report starts with discussing the assessment methodology developed in this task, and then moves to discuss assessment results from applying the proposed methodology on the various network screening approaches. Finally, the report concludes with a summary of the evaluation methodology and a highlight of major findings.

2. ASSESSMENT CRITERIA

This task utilized the criteria developed in Task 3 for assessing the merits and limitations of the network screening methods. While discussing these criteria is considered beyond the scope of this report, a list of the criteria along with brief definitions are provided below.

1. *Sensitivity to Level of Risk*: For the purpose of this project, level of risk is defined as the likelihood of a crash taking place at a site when it is used by motorists. Risk factors are defined as roadway, roadside, traffic, and environmental properties that have a strong correlation with crash occurrence.
2. *Sensitivity to Economic Effectiveness*: Economic effectiveness, often used in the form of benefit-cost ratio, is based on the premise that a site that is expected to yield higher monetary return on safety investment is more deserving to receive safety funds. Sites with higher crash frequencies and higher number of severe crashes tend to yield greater benefits upon implementing safety countermeasures.
3. *Precision*: This criterion is used to assess whether a network screening method can respond to small and subtle changes of any factor related to the level of risk or crash occurrence at a particular site. A less precise screening method might lead to discarding potential at-risk sites, as the method may not be able to accurately assess the risk due to differences in magnitude of a risk-related feature.
4. *Previous Performance Record*: Only when a screening method is applied and validated, a full understanding of the strengths and limitations of that method is achieved. The record of a method being used in practice with satisfactory results is often associated with the practicality of the method and the merits perceived by users.
5. *Ease of Understanding*: This criterion refers to how intuitive or easy to comprehend a network screening method is to the practitioner. Since the prospective screening method will be applied on local and low-volume roads, ease of understanding by staff at state and local agencies is important.
6. *Ease of Implementation*: The ease of implementation, which is closely related to the practicality of a given method, considers factors such as the availability and accessibility of data and whether specially trained personnel are required.
7. *Resource Requirements*: This criterion refers to the resources needed when implementing a prospective network screening method, which primarily involve agency personnel as well as other costs involved in acquiring the data, including staff time.

It is important to note that the data requirements criterion identified in Task 3 is combined with resource requirements for the purpose of the assessment in this task. This is done to simplify the assessment since data acquisition is one aspect of resource requirements.

3. ASSESSMENT OF SCREENING METHODS

The assessment methodology developed and used in this task aims at removing much of the subjectivity of the assessment criteria by following a systematic quantitative approach in expressing the level of a method meeting certain criteria. The method consists of three key elements: developing a scoring scheme for criteria, assigning weights to criteria, and preparing the final assessment matrix. Each of these elements is discussed in detail in the following sections.

3.1. Developing a Scoring Scheme for Assessment Criteria

In this step, numerical scores were assigned to the evaluation criteria based on the level a particular method meets those criteria. The scoring scheme developed is shown in ***Table 1***. A four-level score (1 to 4) was assigned to level of risk representing sites lacking risk factors, those with traffic properties, geometric properties, and those with both traffic and geometric properties involving risk in an ascending order. Cost effectiveness was assigned three levels (1 to 3) with the lowest level for methods overlooking frequency and severity of crashes, middle level for methods considering either frequency or severity of crashes, and the higher level for methods considering both crash frequency and severity in screening the network. Precision was assigned three levels: presence of a feature, a range of values for a feature, and an exact value for that feature with scores increasing with the level of precision. Previous performance record was assigned four levels based on three factors: 1) whether the method was used by agencies, 2) the number of agencies using the method, and 3) any reported evaluation or validation of the respective method. Methods that are proposed in the literature but haven't moved yet into practice are assigned the lowest level. Methods that have been used by one or two highway agencies, but no validation is reported are assigned level 2. Level 3 is assigned for methods used by one or two agencies with reported validation, or methods that are used by three or more agencies without a reported validation. The highest level (level 4) is assigned for methods that are used by three or more agencies with validation or evaluation of the method reported in the literature. Ease of understanding is assigned only two levels: level 1 for complicated methods that may be difficult to understand and level 2 for methods that are intuitive and easy to understand. Ease of implementation is assigned three levels, with the lowest level (level 1) representing methods that are difficult to implement, highest level (level 3) for methods that are easy to implement, and the middle level (level 2) for methods that fall in between the previous two categories. Finally, the resource requirements criterion is assigned two levels, level 1 for methods requiring significant resources and level 2 for methods requiring average resources.

Table 1:Scoring Scheme for Evaluation Criteria

Criteria	Scores	Score Categories	
Level of Risk	1-4	No risk factors	1
		Risk - Traffic properties	2
		Risk - Geometric properties	3
		Risk - Traffic & geometric properties	4
Economic Effectiveness	1-3	No frequency, no severity	1
		Only frequency	2
		Only severity	2
		Crash frequency and severity	3
Precision	1-3	Existence of attributes	1
		Range of values of attributes	2
		Exact values of attributes	3
Previous Performance Record	1-4	Proposed methods	1
		Used by 1 or 2 agencies without any validation	2
		Used by 1 or 2 agencies with validation	3
		Used by 3 or more agencies without validation	3
		Used by 3 or more agencies with validation	4
Ease of Understanding	1-2	Difficult to understand	1
		Easy to understand	2
Ease of Implementation	1-3	Difficult to implement	1
		Somewhat easy to implement	2
		Easy to implement	3
Resource Requirements	1-2	Significant resources	1
		Average resources	2

3.2. Assigning Weights to Assessment Criteria

As it is logical to think that different criteria bear different significance in assessing the merits of screening methods, it was important to develop a quantitative approach where relative weights are assigned to the assessment criteria. This has proven to be a challenging step, yet it could be the most important step in the process. The procedure outlined in this section is intended to quantify a process that is inherently subjective. To simplify comparisons across multiple criteria, the problem was reduced to conducting a large number of pairwise comparisons between pairs of variables using a matrix covering all pair combinations, as shown in ***Table 2***. In this table, green cells are for relative weights assigned to criteria in the column headings, while blue cells are for weights assigned to criteria shown in the row headings. In pair-wise comparisons, weights are assigned to the two variables based on their relative importance (the total weight 10 is split between the two variables, with a minimum weight of

1 and a maximum weight of 9). The weights assigned to each criterion while comparing to all other criteria are then summed to find the total weight which is an indicator of the level of significance of that criterion. Lastly, the total weight is normalized by dividing by the highest possible weight (63 is the highest possible weight for the matrix shown in table 2) and converted to a suitable scale (a scale of 0-10 is used for this assessment).

Table 2:Table Showing the Process of Assigning Weights to Criteria

	Level of Risk	Cost Effectiveness	Precision	Previous Performance Record	Ease of Understanding	Ease of Implementation	Resource Requirements	Total Weight	Normalized Weight (out of 10)
Level of Risk									
Cost Effectiveness									
Precision									
Previous Performance Record									
Ease of Understanding									
Ease of Implementation									
Resource Requirements									
Total Weight									
Normalized Weight									

Note: Green cells show weights assigned to column criteria while blue cells show weights assigned to row criteria.

While the process outlined above largely simplifies the problem at hand (comparison across multiple criteria), pair-wise comparisons could still be difficult and may yield inconsistent results in the absence of any guiding rules. These guiding rules regarding the significance of different criteria are important for accurate and consistent evaluation results. It should be noted that, the guiding rules should reflect, to a large extent, the priorities and perspective of the user, i.e. the agency. In this assessment, the researchers developed the following set of rules to guide the pairwise comparisons in the assessment process.

1. Cost effectiveness is the most important among all other criteria investigated in this project and described earlier. This is in line with the fact that most highway agencies use cost effectiveness in selecting and justifying safety improvement projects.
2. The inherent level of risk to any specific site in the network is the second most important criterion after cost effectiveness and above all other criteria.
3. Previous performance record is more important than other criteria, namely; precision, ease of implementation, ease of understanding and resource requirements.

Using these guiding principles and conducting the respective pairwise comparisons, the overall weights assigned to different criteria are shown in ***Table 3***. The minimum and maximum weights are found to be 4.1 and 8.7 on a scale of 10.

Table 3:Pairwise Comparison Results and the Relative Weights Assigned to Criteria

	Level of Risk		Cost Effectiveness		Precision		Previous Performance Record		Ease of Understanding		Ease of Implementation		Resource Requirements		Total Weight	Normalized Weight (out of 10)
Level of Risk		5		6		2		3		2		2		2	48	7.6
	5		4		8		7		8		8		8			
Cost Effectiveness		4		5		1		2		1		1		1	55	8.7
	6		5		9		8		9		9		9			
Precision		8		9		5		7		5		5		5	26	4.1
	2		1		5		3		5		5		5			
Previous Performance Record		7		8		3		5		3		3		3	38	6.0
	3		2		7		5		7		7		7			
Ease of Understanding		8		9		5		7		5		5		5	26	4.1
	2		1		5		3		5		5		5			
Ease of Implementation		8		9		5		7		5		5		5	26	4.1
	2		1		5		3		5		5		5			
Resource Requirements		8		9		5		7		5		5		5	26	4.1
	2		1		5		3		5		5		5			
Total Weight	**48**		**55**		**26**		**38**		**26**		**26**		**26**			
Normalized Weight	**7.6**		**8.7**		**4.1**		**6.0**		**4.1**		**4.1**		**4.1**			

Note: Green cells show weights assigned to column criteria while blue cells show weights assigned to row criteria.

3.3. Assessment Matrix

The scoring scheme and relative weights developed in the previous steps are used in establishing the assessment matrix and ranking of the network screening methods. The assessment is accomplished using the following steps:

1. All methods are scored using the scoring scheme developed earlier in the process and normalized using a common scale of 0 to 10.
2. The normalized scores are then multiplied by the weights of their respective criteria discussed previously (shown in ***Table 2***) resulting in a weighted score for each assessment criterion.
3. Finally, the weighted scores for the various criteria are summed to yield the overall score for each method included in the assessment.

Table 3 shows the assessment matrix using the steps discussed above.

Table 4:Assessment Matrix for the Different Methods

	Assessment Criteria																												
	(1) Level of Risk				**(2) Economic Effectiveness**				**(3) Precision**				**(4) Previous Performance Record**				**(5) Ease of Understanding**				**(6) Ease of Implementation**				**(7) Resource Requirements**				**Overall Score**
	Actual Score	**Normalized Score**	**Weights**	**Weighted Score**	**Actual Score**	**Normalized Score**	**Weights**	**Weighted Score**	**Actual Score**	**Normalized Score**	**Weights**	**Weighted Score**	**Actual Score**	**Normalized Score**	**Weights**	**Weighted Score**	**Actual Score**	**Normalized Score**	**Weights**	**Weighted Score**	**Actual Score**	**Normalized Score**	**Weights**	**Weighted Score**	**Actual Score**	**Normalized Score**	**Weights**	**Weighted Score**	
	4	**10**	**7.6**		**3**	**10**	**8.7**		**3**	**10**	**4.1**		**4**	**10**	**6**		**2**	**10**	**4.1**		**3**	**10**	**4.1**		**2**	**10**	**4.1**		
Methods Using Only Historical Data																													
Crash Frequency Methods	1	2.5	7.6	19	2	6.7	8.7	58	3	10	4.1	41	2	5	6.0	30	2	10	4.1	41	3	10	4.1	41	2	10	4.1	41	272.49
Crash Frequency and Severity Methods	1	2.5	7.6	19	3	10	8.7	87	3	10	4.1	41	3	7.5	6.0	45	2	10	4.1	41	3	10	4.1	41	2	10	4.1	41	316.67
Crash Rate Methods	2	5	7.6	38	2	6.7	8.7	58	3	10	4.1	41	2	5	6.0	30	2	10	4.1	41	3	10	4.1	41	2	10	4.1	41	291.53
Crash Frequency and Rate Methods	2	5	7.6	38	2	6.7	8.7	58	3	10	4.1	41	3	7.5	6.0	45	2	10	4.1	41	3	10	4.1	41	2	10	4.1	41	306.61
Crash Frequency, Rate and Severity Methods	2	5	7.6	38	3	10	8.7	87	3	10	4.1	41	3	7.5	6.0	45	2	10	4.1	41	2	6.67	4.1	28	2	10	4.1	41	**321.96**
Methods Using Prediction Only																													
Using only Traffic Characteristics	2	5	7.6	38	2	6.7	8.7	58	3	10	4.1	41	1	2.5	6.0	15	2	10	4.1	41	2	6.67	4.1	28	2	10	4.1	41	262.70
Using only Geometric Characteristics	3	7.5	7.6	57	2	6.7	8.7	58	3	10	4.1	41	1	2.5	6.0	15	2	10	4.1	41	2	6.67	4.1	28	2	10	4.1	41	281.75
Using Traffic and Geometric Characteristics	4	10	7.6	76	2	6.7	8.7	58	3	10	4.1	41	2	5	6.0	30	2	10	4.1	41	2	6.67	4.1	28	2	10	4.1	41	315.87
Using Surrogate Safety Measures	2	5	7.6	38	1	3.3	8.7	29	2	6.7	4.1	28	2	5	6.0	30	1	5	4.1	21	1	3.33	4.1	14	1	5	4.1	21	179.89
Comb. of Historical and Prediction Methods																													
Empirical Bayes	4	10	7.6	76	3	10	8.7	87	3	10	4.1	41	4	10	6.0	60	1	5	4.1	21	1	3.33	4.1	14	1	5	4.1	21	**320.11**
Index Methods	4	10	7.6	76	3	10	8.7	87	3	10	4.1	41	3	7.5	6.0	45	1	5	4.1	21	1	3.33	4.1	14	1	5	4.1	21	305.03
Other Methods																													
FHWA Systemic Approach to Safety	3	7.5	7.6	57	1	3.3	8.7	29	2	6.7	4.1	28	4	10	6.0	60	2	10	4.1	41	3	10	4.1	41	2	10	4.1	41	297.88

4. RESULTS

As can be seen in this matrix, the overall scores for different methods roughly ranged between 180 and 320. Only two methods scored greater than 320: methods using crash frequency, rate, and severity (321.96) and the Empirical Bayes method (320.11). These two scores are very close, and for all practical reasons, they can be considered comparable. A careful examination of the assessment matrix clearly reveals that the Empirical Bayes scored very high on all criteria except the last three criteria in the matrix which are assigned lower weights (i.e. less significant criteria). On the other hand, the conventional frequency, rate, and severity methods scored lower on criteria 1 and 4 (level of risk and previous performance record) and higher on the last three criteria. Methods using surrogate safety measures scored lower than all other methods included in the matrix, which is somewhat expected. Another important observation is that the FHWA systemic approach to safety scored right in the middle compared to other methods. Nonetheless, this method could have scored higher than all other methods if the cost effectiveness criterion was assessed more objectively. Specifically, this method scored lowest on cost effectiveness due to the way the scoring scheme was developed (no crash frequency nor severity is used). However, it is well known that systemic improvements consist of low-cost countermeasures that are often associated with relatively high benefit-cost ratios. Taking this into account, it is advised that this method be given serious consideration in later project tasks as its merits were underestimated relative to other methods in this assessment.

5. SUMMARY AND FINDINGS

This interim report presented Task 5 analyses regarding the assessment of the various network screening methods using the set of criteria developed in Task 3 of the project. To a large extent, the analyses aimed to minimize the subjectivity in the assessment process by following a consistent quantitative scoring and ranking techniques. First, a scoring scheme was developed for each of the seven criteria used in this assessment. Next, the seven criteria were assigned relative weights based on their significance in the network screening process. The third and last step involved establishing the assessment matrix using input from the previous two steps and the degree to which each respective method met the assessment criteria. The major findings of the assessment can be summarized in the following:

1. The methods that scored highest in the assessment are the Empirical Bayes method and the conventional frequency, rate, and severity methods.
2. Using surrogate safety measures in network screening scored lowest in the assessment.
3. The FHWA systemic approach to safety scored right in the middle according to the assessment matrix. However, this method was somewhat penalized by the scoring scheme of the cost effectiveness criterion, which relies on the method being sensitive to crash frequency and severity. This method could have scored much higher if the cost effectiveness criterion had taken into account the high benefit-cost ratio often associated with low-cost systemic improvements.

It is important to note that the assessment performed in this task is high level assessment, which did not involve any validation using empirical data. Further, it is also important to mention that the weights assigned to assessment criteria and the level to which those criteria were satisfied by different methods represent the opinions and judgement of the researchers performing the analysis, and not necessarily those of the Montana Department of Transportation.