Safety Improvement Candidate Location (SICL) Methods

By Michael D. Pawlovich, Iowa Department of Transportation, Highway Division, Engineering Bureau, Office of Traffic and Safety

Last Modified: 20. February 2007

Methods for determining Candidate Locations, High Hazard Locations, or Sites With Promise enable practitioners to determine those sites that they focus their limited safety funds on improving [1]. Identification of these locations is a vital component of hazard reduction and safety improvement [1]. Focusing on the locations identified, practitioners can address safety concerns and ultimately reduce crash frequency and/or severity [1].

The federally-mandated Highway Safety Improvement Program (HSIP) required each state to "develop and implement, on a continuing basis, a highway safety improvement program which has the overall objective of reducing the number and severity of crashes and decreasing the potential for crashes on all highways [2]." A comprehensive HSIP consists of three components: planning, implementation, and evaluation [3].

The planning component should consist of processes which [3]:

1. collect and maintain data (including crash, traffic, and roadway data),
2. identify hazardous locations and elements,
3. conduct engineering studies, and
4. establish project priorities (i.e., utilize some type of benefit/cost analysis).

Implementation usually involves taking the results of the last two planning components and defining projects, through design and specification. If these projects meet appropriate funding requirements (including benefit/cost requirements) they will then be constructed or implemented.

Evaluation is performed post-construction or implementation to determine the effectiveness of the projects and to improve future HSIP efforts. Evaluation can many times involve some of the same processes as the planning component, namely data collection, identification, and engineering studies.

The crash or hazard mitigation process, as defined by the HSIP, has sometimes been divided into six steps [4]:

1. identify sites with potential safety problems
2. characterize crash experience
3. characterize field conditions
4. identify contributing factors and appropriate countermeasures
5. assess countermeasures and select most appropriate
6. implement countermeasures and evaluate effectiveness

Step 1 is the same as process 2 of the implementation component, steps 2 through 5 essentially restate processes 3 and 4 from the the planning component, and step 6 restates the implementation and evaluation components. Thus, evidence exists supporting the importance of the identification phase to overall safety improvement efforts, whether they are reactive or proactive. In fact, the identification process is the basis, in both listings, for the further processes, in that identification of sites provides analysts and evaluators with a starting point for further study. Without this, they could potentially be faced with the prospect of analyzing and evaluating innumerable sites.

Given this, the identification process needs to be as accurate and informative as possible, resulting in a defensible listing of the sites that are "most hazardous" or that have the "most promise" of crash frequency and severity reduction. However, creating an accurate and informative identification process is not simple and efforts are ongoing to improve and enhance the identification process with both reactive and proactive purposes in mind. This fits well with the HSIP requirement of continuing development and implementation of a highway safety program.

Current and past methods of determining hazardous locations include the following:

- State-of-the-Practice (those used by public agencies):
    1. Spot Map Method
    2. Crash Frequency/Crash Density Methods
        a. Crash Frequency Method
        b. Crash Density Method
    3. Crash Rate Method
    4. Frequency-Rate Method
    5. Quality Control Methods
        a. Number Quality Control Method
        b. Rate Quality Control Method
    6. Crash Severity Methods
        a. Equivalent Property-Damage-Only (EPDO) Method
        b. Relative Severity Index (RSI) Method
        c. Critical Rate in Combination with Number Criteria
        d. Other Methods
    7. Index Methods
        a. Weighted Rank Method
        b. Crash Probability Index (CPI) Method
        c. Iowa Method
    8. Utilize Complementary Methods for Identifying Hazardous Locations

State-of-the-Practice SICL methods are mainly utilized by public agencies on the state and local levels. Many of them have existed for the past couple decades and have not been updated to reflect recent advances in computing and statistics. However, they perform the base function of an SICL method quite well; they result in a ranking list for consideration by analysts and evaluators.

Spot Map Method [5] - The spot map method involves the creation of a map showing clusters of symbols at spots and on segments of road network. The map is then examined for geographic clustering of crashes and those having the greatest numbers of total crashes (or total crashes of a particular type) are identified as being high crash locations. The spot map method is extremely simple and easy to use, however it only provides a very rough estimate of high crash locations and does not provide a list of such locations. The spot map method is suitable for small areas and low numbers of crashes but fails for large areas or numbers of crashes. In the latter case, another high-crash identification method would be more advisable.

Crash Frequency Method [1, 3, 4, 5, 6] - Closely related to the spot map method, the crash frequency method summarizes the number of crashes for spot locations. Locations are ranked by descending crash frequency and those with more than a predetermined number of crashes are classified as high-crash locations to be further scrutinized for statistical significance.

Application of the crash frequency method involves completion of the following steps for each study location:

1. Determine the crash frequency by computing the annual average number of crashes, preferably for at least the three most recent, consecutive, 12-month periods. Less than three years of data may be used; however, considerable caution must be involved in use of shorter time periods, even for high-volume, high-crash locations.

2. Categorize the location by as many features as reasonable using categories such as:
   a. area type: urban or rural
   b. roadway functional class: arterial, collector, or local (using the higher or highest functional class of the intersecting roadways, where an arterial is the highest class (meant primarily to carry through traffic) and a local is the lowest class (meant primarily to provide access to abutting properties))
   c. number of lanes (the number of through lanes on the widest approach)
   d. predominant traffic control (the presence or absence of signalization)
   e. average daily traffic (ADT) volume (the sum of volumes on all approaches)
3. If previously evaluated locations are being catalogued, insert the new location in its proper order by crash frequency. At a minimum, separate lists for intersections and other spot locations should be maintained. As the list grows, begin to keep lists divided out by more specific combinations of the variables above (e.g., when five or more evaluated locations fall into such a category).
4. Determine the critical crash frequency by using one or both of the following approaches for each location type:
   a. Utilize a list of critical crash frequencies, if one has been developed for your state or region. If none exist, these critical crash frequencies can be computed with crash data for the entire state or region using the following equation:

$$F_{cr} = F_{av} + s_F$$

   where:

   F_{cr} = critical crash frequency,
   F_{av} = average crash frequency for all locations of a given type, and
   s_F = standard deviation of crash frequency for all locations of this type.

   Local critical crash frequencies may also be calculated using this equation and the appropriate statistical methods. That is, if a local critical crash frequency is computed, be sure to verify that the sample size is sufficient.

   b. Choose a number of crashes per year (or per year per mile) which is considered "high" and unlikely to be exceeded by many similar locations. This enables an agency to determine a reasonable number of sites for detailed study. This number is subjective and highly empirical.
5. Compare the location's crash frequency to the critical crash frequency. If the critical crash frequency is equaled or exceeded, classify the location as a high-crash location.

The crash frequency is typically used as a basic measure of the safety at a spot location while crash density is used for roadway sections.

Crash Density Method [1, 3, 4, 5, 6] - Closely related to the crash frequency method, the crash density method summarizes the number of crashes per mile for highway sections. Sections are defined as a minimum length of roadway with consistent characteristics, with the minimum distance used frequently being one mile. Locations are ranked by descending crash density and those with more than a predetermined density of crashes are classified as high-crash locations to be further scrutinized for statistical significance.

Application of the crash density method involves completion of the same steps as for the crash frequency method, but determining crash densities for each study location:

1. Determine the crash density by computing the annual average number of crashes per mile, preferably for at least the three most recent, consecutive, 12-month periods. Less than three

years of data may be used; however, considerable caution must be involved in use of shorter time periods, even for high-volume, high-crash locations. The number of crashes is divided by the segment's length in miles to create a comparison measure with which to rate against other segments.

2. Categorize the location by many features as reasonable using categories such as:
   a. area type: urban or rural
   b. roadway functional class: arterial, collector, or local
   c. number of lanes
   d. predominant traffic control (the speed limit)
   e. average daily traffic (ADT) volume
3. If previously evaluated locations are being catalogued, insert the new location in its proper order by crash density. As the list grows, begin to keep lists divided out by more specific combinations of the variables above (e.g., when five or more evaluated locations fall into such a category).
4. Determine the critical crash frequency by using one or both of the following approaches for each location type:
   a. Utilize a list of critical crash densities, if one has been developed for your state or region. If none exist, these critical crash densities can be computed with crash data for the entire state or region using the following equation:

$$D_{cr} = D_{av} + s_D$$

   where:

   D_{cr} = critical crash density,
   D_{av} = average crash density for all locations of a given type, and
   s_D = standard deviation of crash density for all locations of this type.

   Local critical crash densities may also be calculated using this equation and the appropriate statistical methods. That is, if a local critical crash density is computed, be sure to verify that the sample size is sufficient.

   b. Choose a crash density per year (or per year per mile) which is considered "high" and unlikely to be exceeded by many similar locations. This enables an agency to determine a reasonable number of sites for detailed study. This number is subjective and highly empirical.
5. Compare the location's crash density to the critical crash density. If the critical crash density is equaled or exceeded, classify the location as a high-crash location.

The merits of the crash frequency and crash density methods include their simplicity and the fact that locations with many crashes would be studied. However, no consideration for exposure (e.g., traffic volumes) in the prioritization occurs. This lack can result in misleading results if traffic volumes vary considerably throughout the road system. The crash frequency and crash density methods tend to rank high-volume locations as high-crash locations, even if the relative number of crashes is low given its volume.

Many agencies that use the crash frequency and crash density methods only use them to develop an initial list and evaluate the locations in the list in more detail using other methods.

Crash Rate Method [1, 3, 4, 5, 6] - The crash rate method factors the risk of exposure into the determination of high crash locations. The method uses crash rate (number of crashes divided by vehicle exposure) as a basis for ranking. Rates are given in crashes per million entering vehicles (crashes/MEV) for spot locations and crashes per million vehicle-miles (crashes/MVM) for sections. Locations with higher than a predetermined rate are classified as high-crash locations.

Crash rates are calculated using:

$$\text{Crash rate} = a/v$$

where:

a = the number of crashes at a location during a specified time
v = the traffic volume using the location during that same time

Due to the rarity of crashes, this rate is generally multiplied by one million or one hundred million.

Two kinds of rates are generally computed, one for spots and one for sections:

1. The spot crash rate involves the number of crashes per million vehicles entering the spot:

$$R_i = 2\ (A)\ (1{,}000{,}000)\ /\ (T)\ (V)$$

where:

R_i = spot crash rate expressed in crashes per million entering vehicles
A = number of crashes during the days of the study
T = time period in days
V = total average daily traffic entering and departing the intersection

2. The section rate considers section length in addition to volume. Because road sections vary in length, they provide different exposure to crashes; thus, rates for road sections must be in terms of crashes per one million miles or one hundred million miles. Road sections are generally longer than half a mile and usually 100 million vehicle miles are used. The section rate is calculated using:

$$R_s = (A)\ (100{,}000{,}000)\ /\ (T)\ (V)\ (L)$$

where:

R_s = section rate in crashes per 100 million vehicle miles
V = average annual daily traffic on a section (vehicles per day)
T = period (days) for which crashes are counted, usually 365 days
L = length of section in miles

A stepwise method of determining crash rates and developing a list is as follows:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Calculate crash frequencies at individual spots and crash densities along each established section.
3. Using the section crash rate equation, calculate the crash rate for each established section during the study period.
4. Using the spot crash rate equation, calculate the actual crash rate for each intersection or spot during the study period.
5. For the same period, calculate the system-wide average crash rates for sections and spots. Use the appropriate equation (for sections or for spots), inserting the summation of total crashes, total vehicle miles, and total vehicles, respectively, for each category of location.
6. Select crash rate critical values as criteria for identifying high crash locations. Doubling the system-wide rate is usually reasonable and pragmatic.

Selection of the critical values is not completely necessary. The principal purpose is to limit the high crash location list length in order to expedite investigation. Experience will disclose the proper level for a particular agency. Additionally, an agency might simply consider only a certain number of locations (e.g., the top 200).

7. If actual rates exceed the minimum established criteria, the location is identified as a high crash location and placed on the list for investigation and analysis.

The principle reason for using the crash rate method is that it considers exposure in the form of traffic volume. A road location or section might have a high number of crashes simply due to use rather than its being hazardous. Use of crash rate mitigates this. Generally, the crash rate method provides better results than the crash frequency or crash density methods. However, it is more complex than either of those methods, especially as it adds the further complication of requiring non-crash data.

Use of either the crash frequency, crash density, or crash rate methods to identify hazardous locations has its shortcomings. The two-fold purpose of the crash or hazard mitigation process is identification of unsafe locations and simultaneous designation of areas with greatest promise for crash and/or crash severity reduction. Whereas the crash frequency and crash density methods designate the second purpose and the crash rate method designates the first purpose, neither fully addresses the complementary purpose. Improvement can be achieved through use of the frequency-rate method or the quality control methods. These latter methods are recommended for agencies with large, complex systems.

Frequency-Rate Method [1, 3, 4, 5, 6] - The frequency-rate method is a combination of crash frequency/crash density methods and the crash rate method. Locations are classified as high-crash locations if they have more than the prescribed minimum crash frequency or crash density and higher than the minimum crash rate.

The crash frequency/crash density methods and the crash rate methods have deficiencies that limit their effectiveness. However, if these methods are combined, as they are in the frequency-rate method, it appears possible to eliminate or minimize the effects of the deficiencies.

The steps involved in the frequency-rate method are as follows:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Identify crash frequencies for individual spots and crash densities along each established section.
3. For sections, compute average crash density and crash rates for each category of highway, based on total data for all sections of each category:

$$\text{Average crash density} = E(\text{crash frequency})/E(\text{miles})$$

$$\text{Average crash rate} = E(\text{crash frequency})(10^6)/E(\text{section ADT})(\text{no. of days})(\text{section length})$$

4. For spots, compute average crash frequencies and rates for each category of highway, based on total data for all spots of each category:

$$\text{Average crash frequency} = \text{Total crash frequency}/\text{Total number of locations}$$

$$\text{Average crash rate} = (\text{Total crash frequency})(10^6)/E(\text{location ADT})(\text{no. of days})$$

5. Select critical values for each of the criteria above. Begin by doubling the system-wide average for each highway category.
6. For each section, calculate both the crash density and crash rate.

7. For each spot, calculate both the crash frequency and the crash rate.
8. All locations with crash frequency/crash densities and crash rates both higher than the critical values should be placed on the high crash location lists, one for each category of locations. Comparisons must be made with criteria for the particular category of highway being analyzed.

The crash frequency or crash density is used to create the initial list and the crash rate is used to reorder the final list. The number of sites studied further should be commensurate with the staff assigned to conduct additional studies.

The frequency-rate method combines two methods that have different deficiencies, thus minimizing or eliminating these deficiencies. Sites with high crash frequencies/densities might appear to be problematic but if the traffic volumes are also high, the crash rates might then not be high enough to meet the critical value. On the other hand, sites with high crash rates due to extremely low traffic volumes might have low crash frequencies/densities, thus not meeting the critical values. To be classified as a high crash location, sites must meet both criteria and thus be deemed worthy of additional investigation.

However, in conclusion it must be clarified that the deficiencies might only be minimized. Sites that should be investigated further might not be, resulting in a loss of potential crash reduction. Sites that shouldn't be investigated further might be, utilizing time better spent investigating truly hazardous sites.

Though all of the above methods generate useable lists for hazardous site ranking, none of them include any measure of statistical significance or any statistical control. However, a couple currently utilized methods exist that incorporate some simple statistics, one based on crash frequency/density and one based on crash rate:

Quality Control Methods [1, 4, 5, 6] - Similar to the frequency-rate method, the quality control methods consider various highway categories. These methods assure quality control of the analysis by applying a statistical test for determination of unusual crash frequencies/densities or rates. The analysis involves testing the site crash frequencies/densities or rates against predetermined average values for sites with similar characteristics. The statistical tests are based on the oft-accepted premise that crashes fit the Poisson distribution. The critical values are determined using a function of system-wide average crash frequencies/densities or rates for various highway categories and vehicle exposures at the location being studied (the latter of these for rates only). This function incorporates some statistical control by inserting a Poisson distribution probability constant.

Number Quality Control Method [1, 4] - The number quality control method identifies those sites where crash frequency or crash density is greater or significantly greater than the average crash frequency or density for similar sites across the state or similar region. Similar to the crash frequency and crash density methods, the number quality control method adds some form of statistical control for selecting the critical crash frequency/crash density.

The number quality control method applies a statistical test to determine the significance of a site's crash frequency/density when compared to the mean crash frequency/density for similar sites. The statistical test applied is based on the Poisson distribution, the commonly accepted distribution for crashes. Use of the number quality control method effectively addresses sites with high crash frequencies/densities but low exposures. Inputs for the number quality control method, for identification of hazardous sites, include: average crash frequency/density for site category, crash frequency/density at the site, and level of statistical significance.

Determination of each site category's average crash frequencies/densities must be done with care, considering the nature of the sites and their surrounding environment. Site categorizations must be carefully designated and each site then assigned to a particular category. Site categories can be developed using a variety of features, including: rurality, number of lanes, surrounding land use, road types, etc. The purpose of the categories is to facilitate comparison of site crash frequencies/densities with like sites, to the degree possible.

However, this categorization of sites can be taken to unreasonable limits. Therefore, limiting the number of categories to a number which is tenable (that is, neither too large to be unmanageable or that would reduce sample size below statistical reliability nor too small to adequately describe sites) is strongly advised. One suggested breakdown utilizes a combination of rurality of the roadway (urban or rural) and the number of lanes. The categorization utilized should reflect the question being addressed.

After categories have been established, computation of the average frequencies/densities for each category ensues. Many state transportation agencies calculate statewide averages for many categorizations. To compute the critical crash rate for a site, use the following equation:

$$F_c = F_a + k\,(F_a/M)^{1/2} + 1 / 2M$$

where:

- F_c = the critical crash frequency/density
- F_a = average crash frequency/density for the entire population of sites within the category
- k = a probability constant, where the higher the value of k, the higher the value of the critical crash frequency/density. Some common k values are:
  - k = 3.090 for a 99.9% level of confidence
  - k = 2.576 for a 99.5% level of confidence
  - k = 1.645 for a 95% level of confidence
  - k = 1.282 for a 90% level of confidence
- M = millions of vehicle miles (or kilometers) for sections or millions of vehicles for spots

Use of a high k value will result in a shorter list of critical sites but confidence that those sites are hazardous is increased. Critical crash frequencies/densities for low ADT highways are higher because fewer crashes occur within low exposure sites. Also, the use of multiple years of crash data lowers critical crash frequencies/densities due to the variability of crashes at a site over time.

Using the above equation, develop for each categorization a list of critical sites and order them by a Safety Index, which is simply the actual frequency/density divided by the critical frequency/density. The steps involved in using the number quality control method are:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Compute system-wide average frequencies/densities for each category of highway, based on total data for all sites in each category.
3. For each site, determine the vehicle exposure, M, during the study period.
4. Compute the critical crash rate, F_c, for each site within each category using the equation above.
5. Compute the actual observed crash frequency/density at each site for the same time period.
6. Compare the actual crash frequency/density with the critical frequency/density for each site and prepare a list of all sites within each category with frequencies/densities exceeding the critical value.
7. Compute the Safety Index for each site and rank the list for each category by the Safety Index.

Rate Quality Control Method [1, 4, 5, 6] - The rate quality control method identifies those sites where crash rate is greater or significantly greater than the average crash rate for similar sites across the state or similar region. Similar to the crash rate method, the rate quality control method adds some statistical control for determining the critical crash rate.

The rate quality control method applies a statistical test to determine the significance of a site's crash rate when compared to the mean crash rate for similar sites. The statistical test applied is based on the Poisson distribution, the commonly accepted distribution for crashes. Use of the rate quality control method effectively eliminates sites with high crash rates but low exposures. Inputs for the rate quality control method, for identification of hazardous sites, include: average crash rate (per 100 million vehicle miles) for site category, crash rate at the site, and level of statistical significance.

Determination of each site category's average crash rates must be done with care, considering the nature of the sites and their surrounding environment. Site categorizations must be carefully designated and each site then assigned to a particular category. Site categories can be developed using a variety of features, including: rurality, number of lanes, surrounding land use, road types, etc. The purpose of the categories is to facilitate comparison of site crash rates with like sites, to the degree possible.

However, this categorization of sites can be taken to unreasonable limits. Therefore, limiting the number of categories to a number which is tenable (that is, neither too large to be unmanageable or that would reduce sample size below statistical reliability nor too small to adequately describe sites) is strongly advised. One suggested breakdown utilizes a combination of rurality of the roadway (urban or rural) and the number of lanes. The categorization utilized should reflect the question being addressed.

After categories have been established, computation of the average rates for each category ensues. Many state transportation agencies calculate statewide averages for many categorizations. To compute the critical crash rate for a site, use the following equation:

$$R_c = R_a + k\ (R_a/M)^{1/2} + 1 / 2M$$

where:

R_c = the critical crash rate
with:
Crashes per Million Vehicle Miles (MVM) or Million Vehicle Kilometers (MVK)m used for Sections
Crashes per Million Vehicles (MV) used for spots
R_a = average crash rate for the entire population of sites within the category
k = a probability constant, where the higher the value of k, the higher the value of the critical crash rate. Some common k values are:
k = 3.090 for a 99.9% level of confidence
k = 2.576 for a 99.5% level of confidence
k = 1.645 for a 95% level of confidence
k = 1.282 for a 90% level of confidence
M = millions of vehicle miles (or kilometers) for sections or millions of vehicles for spots

Use of a high k value will result in a shorter list of critical sites but confidence that those sites are hazardous is increased. Critical crash rates for low ADT highways are higher because fewer crashes occur within low exposure sites. Also, the use of multiple years of crash data lowers critical crash rates due to the variability of crashes at a site over time.

Using the above equation, develop for each categorization a list of critical sites and order them by a Safety Index, which is simply the actual rate divided by the critical rate. The steps involved in using the rate quality control method are:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Compute systemwide average number of crashes per MV or MVM for each category of highway, based on total data for all sites of each category.
3. For each site, determine the vehicle exposure, M, during the study period.
4. Compute the critical crash rate, R_c, for each site within each category using the equation above.
5. Compute the actual observed crash rate at each site for the same time period.
6. Compare the actual crash rate with the critical rate for each site and prepare a list of all sites within each category with rates exceeding the critical value.
7. Compute the Safety Index for each site and rank the list for each category by the Safety Index.

As mentioned, the quality control methods utilize a statistical test to refine the decision-making process involved in determining a site's hazardousness. Also, these methods allow agencies to determine priorities by grouping locations according to their functional classification and rank within these classifications. Also, sites having higher crash frequencies than average for their category can be quickly singled out for special attention. Though this improves over the previous methods, it still has noteable deficiencies.

First, the statistical test utilized is somewhat ambiguous and suspect. The addition of the Poisson distribution probability constant adjusts the critical rate equation in order to limit the number of sites judged critical. However, the reasoning behind the use of this probability constant in the equation is somewhat unclear. Adjusting the critical rate by a standard deviation or two fits with standard statistical practice, but the third element in the equation (1 / 2M) has a less clear meaning. Additionally, the entire premise of crashes being distributed as per the Poisson distribution has been questioned in recent literature. The Negative Binomial distribution has, recently, been judged a better representation. This may not matter due to the simplicity of this equation and its intended use, but it might introduce some bias due to overdispersion. Finally, the choice of which k-factor value to pick is highly subjective, giving rise to possible ambiguity in results from year to year.

Second, the method is quite data intensive, if simply because it needs to be in order to achieve the gains. For each site and site category the user must track several different types of data that wouldn't be needed under the spot map, crash frequency/density, crash rate, and frequency-rate methods. The categorization development process involves the subjective determination of categories through examination of site characteristics throughout the jurisdictional region. Many site characteristics are now in computerized databases but not all, thus requiring some data collection. Then, once the site categorizations have been developed, each site must be categorized and the method steps listed above must be run for each site within each category.

Third, only crashes and volumes are included in the equation. While the categorizations address other types of data, as the categorizations become more refined, more data must be collected. Again, this might be the price of better refinement in list generation.

Thus far, none of the methods have addressed the idea of including crash or injury severity into the determination of hazardous site ranking lists. However, there are a series of methods which account for severity for list generation:

Crash Severity Methods [1, 3, 4, 5] - Several methods exist that incorporate severity, either of the crashes or of the injuries, into the SICL process. These methods utilize a variety of methods to incorporate severity measures, including: frequency/density of more severe crashes, rate of more severe crashes, and ratio of more severe crashes. Essentially, those crashes or injuries judged more severe are given more relative weight than those judged less severe. Sometimes the results for each site are then compared, as in the quality control methods, to systemwide averages for similar roadways. This inclusion of severity enables highway agencies to devote more of their safety resources to locations with greater exhibited potential for injury or loss of life, thereby allowing the treatment of these locations for reducing overall system severity.

To define severity of crashes and injuries, a standard definition of severity levels has been defined by the National Safety Council (NSC) and is an American National Standards Institute (ANSI) standard [7]:

Fatal: one or more deaths (commonly signified by K)
A-level injury: incapacitating injury preventing victim from functioning normally (e.g., paralysis, broken/distorted limbs, etc.)
B-level injury: non-incapacitating but visible injury (e.g., abrasions, bruising, swelling, limping, etc.)
C-level injury: probable but not visible injury (e.g., sore/stiff neck)
PDO: property-damage only (commonly signified by 0)

Known as the KABC0 injury scale, it is used commonly in police reporting of crashes. Many of the crash severity methods utilize this scale.

<u>Equivalent Property-Damage-Only (EPDO) Method</u> [1, 3, 4, 5] - In the equivalent property-damage-only (EPDO) method weights fatal and injury crashes against a baseline of property-damage-only crashes. Each of the injury levels (KABC) are given a specific number weight that is compared against property-damage-only crashes, which are given a weight of 1. These weight coefficients are based on the relative average crash costs by severity. K-type and A-type crashes often have the same weight. The weights are incorporated into the SICL process by either computing a EPDO index or an EPDO rate.

The steps involved in utilizing the EPDO method for a site are:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Split the crashes by severity level, assigning each crash to a severity level based on its worst severity injury.
3. Calculate the EPDO Severity Index (SI) using the following equation:

$$SI = [W_K K + W_A A + W_B B + W_C C + P] / T$$

where:

SI = Severity Index for the site
W = the respective weight coefficients
K = frequency of fatal crashes at the site
A = crash frequency involving A-type injuries at the site
B = crash frequency involving B-type injuries at the site
C = crash frequency involving C-type injuries at the site
P = frequency of PDO crashes at the site
T = total crashes at the site

4. Calculate the EPDO index using the following equation:

$$\text{EPDO Index} = W_K K + W_A A + W_B B + W_C C + P$$

where the variables are the same as above.

5. Calculate the EPDO rate using the following equation:

$$\text{EPDO Rate} = [\text{EPDO Index} \times 10^6 \text{ or } 10^8] / [(\text{Exposure per day}) \times \text{Days}]$$

6. Categorize the site as per the quality control methods.
7. Compare the site SI, EPDO Index, and/or EPDO Rate to its respective category critical values to determine the hazardousness of the site. If the site's values exceed the category critical values, include the site on the hazardous site list. Rank the list by either EPDO Index or EPDO Rate.

For step 2 it is important to note that the more severe crash types are less likely to occur. Therefore, several years of data may be required to compute a meaningful EPDO Index or Rate. However, great care should be exercised when using multiple years to insure that traffic and road characteristics have not changed significantly during the analysis period.

The EPDO Method improves on the previous methods in that it includes crash severity. However, the method, like the quality control methods, require more data than the simple crash frequency/density or crash rate methods. Gains in hazardous site identification might be sufficient to warrant this, however.

Relative Severity Index (RSI) Method [3, 4, 5] - The relative severity index (RSI) method incorporates the weighted average cost of crashes at sites. This method is best-suited for the further evaluation of sites already identified by other methods as high-crash sites. In the RSI method, crash frequency at each severity level is multiplied by the average "comprehensive cost" for crashes at that severity level. The subtotals for each of these severity-specific costs are summed and the sum is divided by the total crash frequency.

The RSI method, step-by-step, is:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Split the crashes by severity level, assigning each crash to a severity level based on its worst severity injury.
3. Compute the RSI value for the site, utilizing the severities in the following equation:

$$RSI = [C_F K + C_A A + C_B B + C_C C + C_P P]/(K+A+B+C+P)$$

where:

C_i = the average comprehensive cost per crash for a crash of severity level "i" from K thru P
K, A, B, C, and P are as defined above in the EPDO method.

11. Assign the site into a site category, much like the quality control methods, and compare the site's C_i against the category's critical C_i. If the site has critical C_i, insert it into the list of sites for that category, ranked by C_i.

The RSI method allows for crash severity to be included in SICL list generation. However, it also requires, much like the quality control and EPDO methods, more information about each site than the simpler methods. Additionally, the RSI method, through its use of severity cost values, introduces proxy measures into the computation, rather than utilizing the data as is. If these proxy measures are not accurate, the calculations and lists generated using them will be inaccurate.

<u>Critical Rate in Combination with Number Criteria</u> [1] - The critical rate in combination with number criteria method is based on warrants. The warrants include a concentration criteria and a severity criteria. To meet the concentration criteria, a site has to have exceeded a certain frequency/density of crashes for a period of years and another frequency/density of crashes for one year. To meet the severity criteria, a site must have an EPDO rate exceeding a certain level (e.g., 2 crashes/MEV). Critical rates for total crashes, night crashes, fatal crashes, etc. can also be utilized to determine high-hazard sites.

<u>Other Methods</u> [1] - Some agencies use the ratio of fatal crashes to total crashes. Others calculate fatal crash rates, fatal plus injury crash rates, and total crash rates for each facility type. They then use these average rates to determine a site's hazardousness.

Crash severity methods are an excellent way to incorporate into the SICL process the information that is collected about the cost of crashes to individuals and society. However, not only are they somewhat subjective and thus somewhat subject to error, they also require more data for accurate results. Where crash frequency is small, more severe crashes can quickly control the results even though these more severe crashes might be caused by factors unrelated to the highway condition. If not given proper consideration, the crash severity method results could lead to erroneous expenditures of safety improvement funds for sites where crash severity may not be sensitive to highway treatments. Currently, proper consideration is provided by analysts surveying the crash reports for each of those sites identified as being hazardous. Efforts are underway, however, to automate this process, some effort through database management, some effort through improved, more informative statistical procedures. Another way to mitigate this potentiality is to utilize more information about non-severity indicators in the methodologies.

<u>Index Methods</u> [1, 4, 5] - Three index methods exist which attempt to incorporate severity indices with other previously described methods. These two index methods are the weighted rank method, the crash probability index (CPI) method, and the Iowa Method.

<u>Weighted Rank Method</u> [1, 4] - The weighted rank method combines some of the previous methods in the calculation of a single index value for each site. Many times the weighted rank is created by giving equal weight to as many as five indicators, such as: crash frequency/density, crash rate, percentage of wet crashes, percentage of night crashes, and crash severity (utilizing a simple 5-point scale). A ranked list is prepared for each of the five indicators and then the ranks for each site within these lists are combined based on the weighting schema to produce a combined list. The list thus created is then ranked based on the weighted value.

The premise of the weighted rank method is to retain some benefits from each of the different measures while simultaneously eliminating or minimizing the disadvantages. The method also allows agencies to change weightings based on their priorities. Obviously, using the weighted rank method requires more effort, as an agency is required to produce several lists in order to develop the final weighted list. Also, the weightings determined by the agency, if not carefully researched, can be highly subjective.

<u>Crash Probability Index (CPI) Method</u> [5] - The crash probability index (CPI) method, much like the weighted rank method, combines the results from previous methods: frequency/density, rate, and severity. The combination, in theory, reduces the misleading results for high-volume and low-volume sites while also inserting severity. Again, like the weighted rank method, the CPI method allows analysts to adjust weightings to reflect agency priorities.

As part of the CPI method, when a site has significantly worse than average crash frequency/density, crash rate, or severity distribution, it is assigned penalty points. The overall CPI for a site is a summation of the penalty points across these three measures. A final ranking list for all sites, ranked by descending CPI, is generated.

Application of the CPI method includes:

1. If not already done, locate all crashes in accordance with accepted coding practices.
2. Determine the each site's crash frequency/density, crash rate, and casualty ratio (CR). Utilize the following equation to compute CR:

   $$CR = (F+A+B+C)/(F+A+B+C+P)$$

   where the variables on the right side of the equation are as defined previously in the EPDO method.

3. Categorize the site as per the quality control methods.
4. Determine the critical values for crash frequency/density, crash rate, and casualty ratio. The former two of these are as described in the quality control methods. The critical casualty ratio is determined for the site's category as well.
5. Compute the CPI value for the site by comparing the site's values computed in Step 2 to their critical values as follows:
   a. If neither the crash frequency/density, crash rate, nor the casualty ratio equals or exceeds their corresponding critical values, the CPI for the site is zero.
   b. If the crash frequency/density equals or exceeds the corresponding critical crash frequency/density, assess five penalty points.
   c. If the crash rate equals or exceeds the corresponding critical crash rate, assess five penalty points.
   d. If the casualty ratio equals or exceeds the corresponding critical casualty ratio, assess ten penalty points.
   e. Sum the sub-CPI penalty points to obtain the site CPI.

   To adjust for agency priorities, adjust each of the sub-CPI penalty points appropriately and apply over all sites considered in the same analysis.

6. Remove any zero CPI sites from analysis.
7. Retain sites with non-zero CPIs and classify them as either: first-class (20 points), second-class (10-15 points), or third-class (5 points). First class sites are of highest priority, while third-class sites receive less immediate attention.

The classification point levels should be adjusted if the sub-CPI penalty points have been adjusted.

Again, like the weighted rank method, the CPI method attempts to utilize the best features of the incorporated methods while eliminating or minimizing the bad features. Agency priorities are also accommodated. The CPI method also, however, requires more effort as it incorporates more methods. Additionally, adjustment of the sub-CPI penalty points can be highly subjective.

Old Iowa Method [8, 9] - In Iowa, in an approach similar to that of the Weighted Rank Method, three ranking lists are generated and then the ranks from these three lists are combined into a single rank. The three sub-lists are a frequency rank, a rate rank, and a severity rank, this last based on "value loss" at the site.

The three sub-rankings have historically been generated using a link-node system for crash location. The link-node system involved the placement of nodes at locations including intersections, grade separations, bridges, ramp termini, severe curvature, and railroad crossings. These locations all have a unique identifier for its geographic location. Each crash at these locations is referenced to this unique location, or reference node. Crashes between these locations are referenced to both the nearest node (the reference node) and the node at the other end of the roadway link (the direction node), with a distance from the reference node specified as well. The total number of crashes that occur at each reference node and reference node/direction node pair can then be easily tabulated. However, only a list for reference node crashes is generated. To enter the first list the number of crashes must meet one of three certain criteria: a fatality, X number of injury crashes, or Y number of property damage crashes. Currently, X is set at 5 and Y is set at 8. This list typically results in 10,000 to 11,000 locations annually. However, the link-node system has been abolished and a switch to a coordinate-based system is in effect. Adjusting the Iowa SICL method to reflect this is one of the challenges for the Office of Traffic and Safety.

The first two rankings lists are generated much the same as, respectively, the crash frequency/density methods and the crash rate method. Because Iowa has historically relied on a link-node system, the definition of a site, whether spot or section, is slightly affected. In fact, three different types of sites were generally defined:

1. **Intersections** include all road-to-road intersections, except alleys, ramp terminals, and complex intersection or interchange sites.
2. **Links** include sections of road between intersections or nodes.
3. **Nodes** include rail to road intersections, grade separations, bridges, road ends, 90 degree turns, county lines, and major signalized commercial entrances.

Steps involved in the Iowa Safety Improvement Candidate Location (SICL) development process are:

1. The crash statistics are searched to identify all locations (intersections, links, and nodes) in the State that meet at least one of the following crash frequency requirements for the designated five-year time period to develop the candidate location file:
   a. at least one fatal crash, or
   b. at least four personal injury crashes, or
   c. at least eight total crashes.
2. The candidate location file created in Step 1 is sorted by descending frequency of crashes and a frequency rank is assigned.
3. For each site in the candidate location file, the frequency of each category (as defined by the KABCO scale) of injury is determined. A value loss is determined using these injury severity frequencies using the following values (updated in 2001):

   a. Fatalities x $1,000,000, plus
   b. Major Injuries x $150,000, plus
   c. Minor Injuries x $10,000, plus
   d. Possible/Unknown Injuries x $2,500, plus
   e. Actual Total Property Damage or $2,500 if unknown.

   A value loss rank, generated by sorting the value losses in descending order, is assigned.

4. Crash rates per million entering vehicles are calculated for sites with known traffic exposure data. The sites are sorted by rank in descending order and a crash rate ranking is assigned to each site. Sites with no traffic exposure data are initially assigned a rank of 0 to give these sites the highest possible priority in the rate ranking. Traffic volumes are then determined, from any credible source, for sites with a rate rank of 0 that fall within the top 200. This process continues until all sites within the top 200 have valid rank values for rate.

   Crash rates per million entering vehicles are calculated as:

   Rate=(Frequency)(1,000,000)/(DEV)(5 Years)(365 Days/Year) where DEV is the actual DEV for spot locations and road segments up to 0.6 miles long.

For road segments 0.6 miles long and longer the DEV is calculated as:

DEV=ABS((Link Length/0.3)(DEV))

This calculation adjusts the daily entering vehicles by the number of 0.3 mile sections within the segment to correlate the crash rate for longer segments closer to that for a spot location or shorter segment. This is an attempt to enable comparisons between spot locations and segments and enables one rank list, rather than 2 or 3, to exist.

5. The three rankings, frequency, value loss, and rate, are summed to create a composite rank factor. The sites are then sorted in ascending order by this composite rank factor and assigned a composite state ranking.

The Iowa method has many of the same positive features and negative features of those methods it incorporates: frequency, rate, and severity.

New Iowa Method (Intersections) - In Iowa, the approach used is similar to that of the Weighted Rank Method. Three ranking lists are generated and these three rank lists are subsequently combined into a single rank. The three sub-lists are a frequency rank (total crashes), a rate rank (crashes/volume), and a severity rank ("value loss" at the site).

The first step in the process is to identify the crashes that can be assigned, for this purpose, to each intersection. The crashes within 75 feet of urban intersections and 150 feet of rural intersections are assigned to the intersection, using a Geographic Information System (GIS). This information is then exported to a file which is later imported into SAS (SAS Institute Inc., Cary, NC). The file contains the crash assignation data for all intersections that have had at least one crash in a five-year time range. There are roughly 45,000 such intersections that meet this criteria.

Within SAS, the values for the separate rankings as well as the combined ranking are calculated. The first two ranking lists (frequency and rate) are generated much the same as the crash frequency/density methods and the crash rate method. The third ranking list (severity) is generated using a severity index method, based on criteria determined by the Iowa Department of Transportation (Iowa DOT) Office of Traffic and Safety (TAS). The three are combined using a weighting method, determined by TAS, to emphasize high severity locations.

Steps involved in the Iowa Safety Improvement Candidate Location (SICL) development process are:

1. The crash statistics are searched to identify all locations (intersections) in the State that have, for the designated five-year time span, at least one crash. There are typically roughly 45,000 intersections that meet these criteria. A file identifying cases assigned to each intersection is generated. A file detailing the road segments entering each intersection is also generated.
2. Both files are imported into SAS and further analyses are performed:
   a. The crash frequencies for the five-year time span are calculated. The frequencies determined include total crashes, total fatalities, and total major injuries.
   b. The daily entering vehicles (DEV) for each intersection are calculated by summing the 2-way volumes for each road segment associated with each intersection and dividing by 2. This is not absolutely correct given the nature of the road segmentation but it is a compromise made due to the systematic, statewide nature of the analyses and the large number of intersections for which data needs to be obtained. (An Iowa intersection database is under development.)

c. Given the total crash frequencies and the DEV, the crash rates are calculated.
d. Given the injury severity level frequencies, the severity indices are calculated using the following procedure:
    i. The first fatality at any one site is converted to a major injury to partially mitigate the effect of random chance, seatbelt use, age-related skeletal-musculature frailty, etc.
    ii. The following values are multiplied against frequency of injury severity level:
        1. Fatality → 200
        2. Major Injury → 100
        3. Minor Injury → 10
        4. Possible or Unknown Injury → 1
    iii. These values are summed for each intersection to determine the severity indices.
e. Each category (frequency, rate, and severity) are ranked individually. Ties are allowed.
f. The rank list for each category is normalized using the highest rank value. Thus, if the highest rank value for rate is 5,000, all rank values for rate are divided by 5,000. The normalization is done to minimize the impact of any large number effect within a particular rank list when calculating the combined value for the subsequent combined rank.
g. The three normalized rank lists are weighted using values of 1/5 for frequency, 1/5 for rate, and 3/5 for severity index. The combined value is attained by summing these three.
h. The combined values are used to produce the combined statewide rank list. This list with a host of supporting information is exported to file.

3. Within Excel, column headers, borders, headers, and the like are applied to the list table.

Though all these methods develop lists for further consideration, they are not the only ways that sites can be identified as hazardous. Many non-crash based methods exist which might aid in proactively determining hazardous locations prior to existence of a crash history. These methods may also complement the identification of hazardous sites by verifying the existence of problems or by clarifying those problems.

Utilize Complementary Methods for Identifying Hazardous Locations [1] - Complementary methods utilize non-crash indicators to aid in identifying the most hazardous location. They include:

1. Results of road skid testing
2. Hazard Indicator reporting
3. Observed minor crashes
4. Observed near-crashes
5. Evidence of potential hazards such as skidmarks at intersection approaches
6. Maintenance records
7. Median or shoulder encroachment wheel marks
8. Volume to capacity ratios
9. Stopping and passing sight distance
10. Access points (driveways)
11. Traffic conflicts analysis
12. Erratic maneuver observations
13. Reports of hazardous locations by highway personnel, police, department personnel, motor clubs, motorists, and others.

Though all of these "state-of-the-practice" methods have proven useful, none address the identification of high crash locations thoroughly. In addition to the problems with each stated previously, all the methods ignore a significant majority of the system-wide sites in their analyses. Sites without any crashes in the time period analyzed are routinely ignored. This directs all mitigation measures to a reactive, rather than proactive, role. While consideration of only those sites having a crash history makes direct sense from a crash reduction standpoint, consideration of sites without a crash history is more difficult to justify. However, inclusion of sites without a crash history allows for analysis of those factors about the sites that might lend themselves to safety or the lack thereof. Of course, to determine the problems on a systematic basis requires much more effort than obtaining crash histories and traffic volume data. To properly analyze sites to determine their deficiencies, a system-wide database containing the relevant attributes must be polled, thereby increasing the level of effort required to create a ranking list.

References

1. Program of Instruction for the Workshop on Identification and Treatment of High Hazard Locations. The Traffic Institute, Northwestern University, Evanston, IL. November 3-5, 1999.
2. FHWA, "Federal-Aid Highway Program Manual", Volume 8, Chapter 2, Section 3, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C. (March 5, 1979).
3. Methods for Identifying Hazardous Highway Elements. National Cooperative Highway Research Program Report 128. TRB/National Research Council, Washington, D.C., 1986, 80p.
4. Accident Mitigation Guide for Congested Rural Two-Lane Highways. National Cooperative Highway Research Program Report 440. TRB/National Research Council, Washington, D.C., 2000, 164p.
5. SEMCOG Traffic Safety Manual (Second Edition). Southeast Michigan Council of Governments, Detroit, MI, September 1997.
6. Traffic Safety Fundamentals Handbook. Prepared by Howard R. Green Company for Minnesota Department of Transportation, Office of Traffic Engineering, April 2001.
7. Ogden, K.W. Safer Roads: A Guide to Road Safety Engineering. Avebury Technical, Aldershot, England, 1996.
8. Estochen, Bradley Michael. An Evaluation of Medical Crash Costs: A CODES Application. Thesis. Iowa State University, 1999, 85p.
9. Iowa Department of Transportation Office of Traffic and Safety (TAS) writeup of the Iowa Safety Improvement Candidate Location (SICL) methodology.