DATA COLLECTION AND IN-SERVICE EVALUATION ISSUES

William W. Hunter

INTRODUCTION

When considering the elements of a roadside safety management system, appropriate data about system components are an important ingredient. Only with such data can in-service evaluation be performed. This paper is not meant to be an exhaustive primer on data collection and evaluation for a roadside safety management system, but rather an overview of selected relevant issues.

An important underlying point is that while appropriate data elements are crucial to any attempt at evaluation, data collection by nature can be time consuming and expensive. Thus, analysts or evaluators need to decide what will be studied and for what duration, and then identify necessary data. An effective roadside safety management system does not have to include every piece of conceivable data about every feature.

Another important point of departure is the great need to perform more in-service evaluations of roadside safety improvements and to publish the results. The following statement, made fifteen years ago, still has relevance (Hunter et al., 1977):

> "... The fact that the estimates of effectiveness are not more specifically defined is a major roadway safety issue. There is a continuing very serious need for more well-designed effectiveness evaluations of fixed object treatments . . . there is a scarcity of good evaluations concerning fixed object improvement programs. Where such evaluations exist, they generally are the before/after type with no control group and thus are subject to accident fluctuations, regression to the mean, and other artifacts."

With continual changes in the vehicle fleet, continuing evaluations of roadside features are necessary. For example, how well do current guardrail designs perform? There are many guardrail and median barrier end treatments in existence - how well are they working? Which side slopes are safe and which produce rollover? What can we say today about warrants for guardrail and median barrier? Under what conditions should substandard guardrail be upgraded?

INHERENT PROBLEMS WITH CRASH EVALUATIONS

It is assumed that any modern day roadside safety management system links data about the features with associated crashes. This is because evaluations based on crash experience are still the most readily accepted by decision makers. However, evaluations based on proxy or surrogate measures that are directly related to crashes are certainly acceptable.

Part of the difficulty with evaluations relates to the use of crashes as the main dependent variable. Crashes are basically rare events. When coupled with the fact that individual treatments usually have a relatively low overall level of effectiveness (e.g., 20 percent accident reduction) and thus can realistically reduce only some proportion of the crashes that occur, it is the case that much data are needed to accurately evaluate treatments. Other limitations of crash data have been noted by Zegeer (1982) and reported as the following in NCHRP Synthesis 128:

- Large variations exist among agencies in accident reporting thresholds (the minimum amount of vehicle damage, represented in dollars, that must occur before an accident is reported by police). This results in many property-damage-only accidents not entering the system. Therefore, these accidents are underrepresented in the accident data base.
- Data errors and inconsistencies are sometimes found within the accident records themselves. For example, contributing circumstances, such as the presence of alcohol and other factors, may or may not be known by the officer who arrives on the scene of the accident.
- Accurate location of accidents, particularly in rural areas, varies widely among state and local agencies. The location referencing methods may not be used properly. Also, many accidents are never reported to a specific location and, therefore, cannot be tied to a known point along the roadway.
- In some states, it may take a year or longer for accident data to become available for analysis.
- Many accidents are not caused by problems with the roadway environment, but may result from vehicle

malfunction (e.g., brake failure), driver error (e.g., drunk driving), or severe weather (e.g., heavy fog, icy roads). These accidents may not indicate a treatable safety hazard at a given location.

In addition, locations not yet defined as high accident locations may have high potential for such a rating. These are sometimes defined by roadside rating schemes. Several factors affect the degree of hazard for roadside objects (Zegeer, 1986), including:

- Object type and rigidity,
- Distance to the object,
- Traffic volume (exposure), and
- Other interacting roadway or roadside characteristics, such as type and condition of shoulder and side slope, curvature, pavement width, cross slope, etc.

A system wide correction of a common problem may then yield good cost-effectiveness. Examples include tree removal from near the roadway and side slope improvements. The important point is that high accident location lists should not stand alone but be combined with inventories of roadside hazards.

Building the Data Base

Many state and local agencies already have in place inventories that would be a natural part of a roadside safety management system. These include information about roadside objects, guardrail and median barrier, bridge structures, roadway features, etc. The normal methods of building these inventories are manual surveys, photologging, videologging, and automated or semi-automated devices (Zegeer, 1986).

Manual Methods

Manual methods involve individuals traveling the road with a vehicle that measures distance, while filling in data collection forms along the way. Tape measures and measuring wheels may be used periodically, to obtain information about side slopes or other objects that are difficult to view through photo - or videologs. Using maintenance crews to visually inspect for run-off-road items like guardrail damage (on length of need or ends), median barrier displacement or scuff marks, crash cushion damage, sign support damage, etc. can result in useful data items. A worthwhile contribution to the literature was made in such a study from New York State (Bryden and Fortuniewicz, 1986) where field investigators followed up at 3,302 barrier crash sites. Data gathered included vehicle size and type, barrier type and rail height, and other highway measures. Performance measures were vehicle containment, secondary collisions, and resulting occupant injuries.

Photologging

The photologging technique uses photographs along a roadway section at equal intervals, such as 50 or 100 feet, and can be used to document roadside safety features. Grids or reference lines are used to obtain lateral placement or heights of roadside objects. Laser video discs may now be used for storage in lieu of film.

Videologging

Videologging is similar to photologging except that a video camera and recorder replace standard photographic methods. With audio, additional information can be stored orally.

Automated and Semi-Automated Methods

These methods include microcomputer recording systems, pavement inventory systems, and change sensing systems and tend to supplement the techniques described previously. Elements that could be obtained this way include (Datta, et al., 1985):

- Degree of curvature (horizontal alignment);
- Vehicle speed;
- Friction factors and skid resistance;
- Roadway roughness;
- Vertical alignment and clearance;
- Elevation and grade;
- Cross-section features;
- Pavement and subsurface conditions;
- Existence of underground utilities; and
- Roadside or right-of-way features.

Weighting Schemes

Many agencies that inventory roadside objects also use a weighting scheme to enable a determination of degree of hazardousness associated with spots, sections, or corridors. Oakland County in Michigan has been progressive in regard to data base activities and uses weights for the following criteria (Zegeer, 1986):

- Whether or not the roadway is curbed.
- The presence of horizontal curves (inside or outside).
- The presence of vertical curves (positive or negative grade).
- The rigidity of the object.
- Average daily traffic (ADT).
- Speed limit.
- Distance from pavement edge.
- Roadway type (county primary, local, etc.).

Severity factors are also assigned to fixed objects, and these are based on a publication by Horodniceanu and Cantilli (1979):

FIXED OBJECT	FACTOR
Utility Poles (wood)	4
Supports - Rigid (steel)	4
Supports - Breakaway	1
Guardrail	3
Bridge Abutment/Wall Face	3
Bridge Abutment & Pier End	5
Bridge Rail Faces	1
GM Barrier	1
Bridge Rail End	5
Fill Slopes	
2:1	5
3:1	4
4:1	3
5:1	2
6:1	1
Cut Slopes	
0.5:1 - 1:1	5
1.5:1	4
2:1	3
3:1	2
4:1 or flatter	1
Hydrant	3
Signposts	1
Trees (diameter)	
Greater than 13"	5
11" - 12"	4
8" - 10"	3
5" - 7"	2
2" - 4"	1
Rocks & Boulders (diameter)	
Greater than 3'	5
2' - 3'	4
1' - 2'	2
Less than 1'	1
Steel beams, concrete posts, etc.	3
Wood posts	
8" x 8"	2
6" x 6"	2
4" x 4"	1
Guy Wire	3
Wood Posts	
6 x 8" guardrail	2
7" round marker post	2

The weighting factors and fixed object severity factors are then combined to form a priority factor, or a measure of relative hazard.

Another example of a weighting scheme used in a roadside safety management system is a recent TRB paper that describes guidelines for the installation of guardrail in the State of Kentucky (Pigman and Agent, 1991). Since guardrail placement was the focus, a list of locations was generated with critical rates of run-off-road accidents. A hazard-index point system was then devised which included characteristics pertaining to both accidents and accident potential, as shown below:

Characteristics	Rating Points Possible
1. Number of run-off-road accidents	15
2. Run-off-road accident rate	15
3. Traffic volume	10
4. Speed limit or prevailing speed	10
5. Lane and shoulder width	10
6. Roadside recovery distance	10
7. Embankment slope	10
8. Embankment height	10
9. Culvert presence	5
10. Subjective roadside hazard rating	5

The authors offered further elaboration:

"An attempt was made to include character- istics representative of accidents and accident potential, operations, and cross section. Point-system weightings of each characteristic were determined by subjective evaluation. The result was combining number of accidents and accident rate to make up 30 of a possible 100 points. Traffic volume and speed limit, considered to be operational characteristics, totaled 20 of the possible 100 points. Cross-section characteristics made up an additional

40 points. Because of their frequency of occurrence and the hazard associated with culvert headwalls or openings near the roadway, a special category was created to represent this condition. or a culvert present within 5 feet of the road, 5 points were assigned. Also included was a general category representing a subjective roadside hazard rating with 5 points possible. This rating was based on a visual observation that was compared with photographic documentation of roadway sections depicting various degrees of roadside hazard."

The end product was the tabulation of hazard-index points after a field survey enabled values to be assigned to the ten factors listed above. This led to various lists of locations rank ordered by total hazard-index points. Costs and benefits were then determined and a budget optimization procedure applied to yield the locations for guardrail placement.

Roadside Rating Scale

In a research study performed for FHWA in the mid-1980's that pertained to the safety of cross-section design on two lane roads (Zegeer, Hummer, Reinfurt, Herf, and Hunter, 1987), a pictorial roadside rating scale was developed by the Highway Safety Research Center. Some 13 highway and roadside safety professionals participated in a workshop and viewed several hundred photographs of roadside situations from both rural and urban two-lane roadways. The situations were then rated in three ways: (1) potential frequency of accidents, (2) potential severity of accidents, and (3) overall hazard. Eventually three 7-point ordinal scales were developed as related to frequency, severity, and overall hazard (which was actually a combination of frequency and severity). Participants used the three scales to rate 141 rural and 78 urban photographs of roadway/roadside situations. Descriptive statistics were then examined to determine which scale produced the most consistent ratings. The overall hazard scale was clearly superior to the other scales, and a 7-point rural and urban scale was then utilized in the study. Examples of the rural scale are shown in Appendix A.

The hazard scales were then used to rate rural and urban roadway sections every tenth of a mile within the data base. As the statistical analysis progressed, the roadside rating scale was reliable and significant enough to be used as a variable in models developed to depict single vehicle accident rates for the rural sections. It was found that a reduction in roadside hazard rating of 1 (e.g., from 7 to 6, or 6 to 5, etc.) due to a roadside improvement would be expected to reduce related accidents by 19 percent. Similarly, a reduction in roadside hazard of 5 (e.g. from 7 to 2) would be expected to reduce related accidents by 65 percent. The point made here is that a pictorial rating scale may have a place in the development of a roadside safety management system.

IN-SERVICE EVALUATION

Background

As mentioned in the introduction, there remains a continuing need for well conducted in-service evaluations of roadside safety improvements. The typical cycle is one of collecting data, identifying problems, examining treatment alternatives, implementing a treatment, evaluating the effect of the treatment, and feeding back what was learned into one or more of these cycle elements. Evaluation remains so important because there will always be limited resources available to treat all the possible safety problems in any given jurisdiction. Thus, prioritizing of projects can be done by examining the results of evaluations.

"Bottom Line" Measure - Number of Crashes or Crash Severity?

It may seem rather simplistic, but simply deciding what to measure has been a problem in many published evaluations. Before beginning any evaluation, an appropriate question is, "What is the treatment supposed to accomplish?" The evaluator should seek to define the answer as narrowly as possible, focusing on the types of crashes that can be affected by the treatment, so that other confounding variables may be limited. As examples:

- Signs warning of hazardous curves should affect the number of crashes on the curves.
- Railroad grade-crossing devices should reduce the number of motor vehicle-train collisions.
- Application of median barrier should reduce the number of cross-median, head-on crashes.
- Placement of a crash cushion in a hazardous gore area should reduce the severity of injuries to the occupants involved in subsequent collisions.

There are any number of examples that could be listed. Besides defining the expected treatment outcome as narrowly as possible, the evaluator must also decide whether frequency of crashes or crash severity is the most important indicator of success. Using an example from above, it is quite likely that number of accidents in a hazardous gore could increase after placement of a crash cushion, simply because there may be less space to maneuver. However, occupant injury severity should decline when striking the crash cushion as opposed to the sign support, bridge end, etc. that the attenuator is shielding.

Threats to Validity

There are many threats to the validity of evaluations, including other things taking place at the same time (history), trends over time (maturation), regression to the mean, and data instability (see Campbell and Stanley (1963) and Campbell (1975) for classic papers). While all of these threats tend to be present in highway safety evaluations, the problem of regression to the mean has likely led to the most erroneous conclusions (Council et al., 1980). The situation tends to occur when sites having the worst recent accident histories are chosen for treatments. The accident problem is many times simply related to fluctuation, and the accident situation likely would have improved (or regressed toward the mean number of accidents) without any treatment at all.

Common Evaluation Designs

Probably the most common design in highway safety has been the simple before-after design, where results are stated based on measurement before and after a treatment is applied (often at a single site). The assumption is made that the after measurement would be similar to the before without any intervention or treatment. There is nothing in this design which controls for the major threats to validity.

A much stronger design would be a before/after with randomized control groups, where candidate sites for a treatment are randomly assigned to either a treatment or control group. Here the predicted after experience is based on the experience of the control groups. Thus, other factors that change simultaneously with the treatment can be accounted for in the results.

Highway departments tend not to follow this procedure because:

1. Treatments are often implemented across an entire jurisdiction; or
2. Engineers think it is ethically or morally wrong to identify hazardous locations and not apply a treatment.

The main rebuttals to these factors are that there are never enough funds to handle all hazardous locations, and treatments at candidate sites can generally never be all made at the same time (Council et al., 1980). Thus, some room is left for evaluation.

A less strong but effective design that is often more acceptable to engineers is the before-after with a non-randomized comparison site. In other words, a comparison site is chosen (sometimes after a treatment is implemented) that is similar to the treatment site. The key to this design is to match the treatment and comparison locations as closely as possible, so that they tend to behave similarly except for the intervention.

Certainly there are other evaluation designs that may be chosen depending on the situation. These include interrupted time series, time series with comparison groups or variables, modeling, etc. (Council et al., 1980). Staff at the Texas Transportation Institute have been active in recent years in developing spin-offs to traditional designs. For example, a 1992 publication by Griffin discusses the use of multiple treatment and comparison groups in before-after designs. (For other examples see Griffin, 1990a; Griffin, 1989a; Griffin, 1989b; Griffin, 1989c.) There are also several excellent evaluation primers available (Zegeer, 1981; Council et al, 1980; Griffin, 1990 b).

Suggested Data Elements and Studies

Given that there are roadside safety management systems already in place in a variety of jurisdictions and likely more to emerge in the near future, it is instructive to think about ways these systems might be used to fill gaps in knowledge. A few suggested data elements and studies follow, but there are many other ideas that could be developed and pursued.

Recently Hunter, Stewart, and Council (in press) utilized the data collected from the Longitudinal Barrier Special Study (LBSS) in the early 1980's to examine driver injury for various guardrail and median barrier designs, as well as crashes into ends versus length of need. A variety of data was collected by trained investigators, including vehicle speed, angle of impact, yawing angle, barrier performance, subsequent vehicle trajectory, etc. These types of data elements would be useful to have in a data system, but it is acknowledged that the items are difficult and expensive to collect. However, more information about real world barrier crashes is necessary as vehicle designs continually change, so that hardware can be compatible. It is also worth noting that much more dialogue should take place between vehicle designers and roadway design engineers, so that perhaps the vehicle design changes that are made (e.g., bumper height) are done so with the knowledge of what hardware will be in existence and the safety consequences of the vehicle change.

Thus for barrier crashes, useful data elements include:

- Vehicle impact speed and angle;
- Vehicle yawing angle, or some measure of whether the vehicle was tracking;
- Barrier impact point;
- Barrier performance (e.g., redirected, penetrated, vaulted, etc.);
- Subsequent vehicle trajectory (e.g., remained on roadside, returned to roadway, crossed roadway, rolled over, etc.); and
- Barrier system type - old or new? Designed or installed properly? Correct height?.

The study referenced above concludes that crashes into ends have worse safety consequences than crashes into length of need, but the findings are somewhat tenuous. With more commercial end treatments coming into existence, we need to know more about how well they work and how they compare to other systems. Along the same line, knowledge about the performance of the breakaway cable terminal (BCT) would be valuable. There is also a need for warrants concerning when to upgrade guardrail. An older system may still be quite functional for a lower service level roadway, but engineers need more guidance about what road and traffic conditions require higher performance systems. Studies concerned with defining crash severity data are also recommended. In other words, what crash severity is associated with striking various fixed objects? How does crash severity change when speed, angle, and other variables are controlled for? These answers have direct relevance to benefit - cost algorithms, and can also be used in roadside hazard weighting schemes.

Besides barriers, numerous other roadside safety features can be tracked, including attenuators, breakaway supports, luminaries, trees, side-slopes, etc. The objective of data collection should be to help understand the nature of the crash and the cause of injury.

Engineers working with a roadside safety management system should be continually thinking about the role that maintenance staff can play in either data collection or input to overall safety feature performance. These people are viewing the roads and roadsides on a daily basis and may be able to provide much insight into problems. For example, if run-off-road crashes are occurring frequently, can maintenance staff offer any reasons why this is happening? Is the problem one of skid resistance? Roadway or roadside cross section? Shoulder problem? Etc. Maintenance personnel should also be trained to understand why systems are designed certain ways, so that repair after crashes is done properly.

RECOMMENDATIONS

To conclude, a few overall recommendations are offered:

1. Data collection for a roadside safety management system can be time consuming and expensive. Where funding is a consideration, think of the data elements needed to answer the questions of interest. It is not necessary to be "all encompassing" every year.
2. To make the best use of a roadside safety management system, seek out the aid of a statistician or an analyst well versed in the proper use of statistics. These persons can help ensure proper experimental designs, as well as extract the most from the data.
3. Publish the results of studies. Too often good studies are done by highway departments or local engineers that are not published in any way. These efforts can advance the state-of-the-art if made available to others.

REFERENCES

Hunter, W.W., Council, F.M., and Dutt, A.K. *Project selection for roadside hazards elimination*. Final report (Vol. I). Chapel Hill, NC: University of North Carolina Highway Safety Research Center, 1977.

Zegeer, C.V. *NCHRP Synthesis of Highway Practice 91: Highway accident analysis systems*. Washington, D.C.: Transportation Research Board, 1982.

Zegeer, C.V. *NCHRP Synthesis of Highway Practice 128: Methods for identifying hazardous highway elements*. Washington, D.C.: Transportation Research Board, 1986.

Datta, T.K., Tsuchiyama, K.H., and Opiela, K.S. *Cost-effective inventory procedures for highway data: User's manual* (Draft report). Washington, D.C.: Federal Highway Administration, 1985.

Horodniceanu, M. and Cantilli, E.J. *Transportation system safety*. Lexington, Massachusetts: Lexington Books, 1982.

Pigman, J.G. and Agent, K.R. *Guidelines for installation of guardrail*. Transportation Research Record 1302. Washington, D.C.: Transportation Research Board, 1991.

Zegeer, C.V., Hummer, J., Reinfurt, D.W., Herf, L., and Hunter, W.W. *Safety effects of cross-section design for* *two-lane roads. Volume I, Final Report*. Washington, D.C.: Federal Highway Administration, 1987.

Campbell, D.T. and Stanley, J.C. *Experimental and quasi-experimental designs for research and teaching*. In N.L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Campbell, D.T. *Reforms as experiments*. In E.L. Struening and M. Guttentag (Eds.), Handbook on evaluation research. Beverly Hills, California: Sage Publications, 1975.

Council, F.M., Reinfurt, D.W., Campbell, B.J., Roediger, F.L., Carroll, C.L., Dutt, A.K. and Dunham, J.R. *Accident research manual*. Washington, D.C.: Federal Highway Administration, 1980.

Griffin, L.I. *Using the before-and-after design with multiple treatment and comparison groups (and multiple before-and-after measurements) to estimate the effectiveness of accident countermeasures*. College Station, Texas: Texas Transportation Institute, 1992.

Griffin, L.I. *Using the before-and-after design with yoked comparisons to estimate the effectiveness of accident countermeasures implemented at multiple treatment locations*. College Station, Texas: Texas Transportation Institute, 1990a.

Griffin, L.I. *Using before-and-after data to estimate the effectiveness of accident countermeasures implemented at several treatment sites*. College Station, Texas: Texas Transportation Institute, 1989a.

Griffin, L.I. *A comparison of two test statistics used to evaluate highway safety improvement projects: A simulation study with samples drawn from Poisson populations*. College Station, Texas: Texas Transportation Institute, 1989b.

Griffin, L.I. *A systematic framework for analyzing categorical, before-and-after data*. College Station, Texas: Texas Transportation Institute, 1989c.

Zegeer, C.V. *Highway safety improvement program - User's guide*. Washington, D.C.: Federal Highway Administration, 1981.

Griffin, L.I. *Engineer's guide to program and product evaluation*. Washington, D.C.: Federal Highway Administration, 1990b.

Hunter, W.W., Stewart, J.R., and Council, F.M. *An analysis using the longitudinal barrier special study file: Comparative performance of barrier and end treatment types* (contract DTFH 61-89-C-00015). Washington, D.C.: Federal Highway Administration, in press.

APPENDIX A
EXAMPLES OF RURAL ROADSIDE RATING SCALE

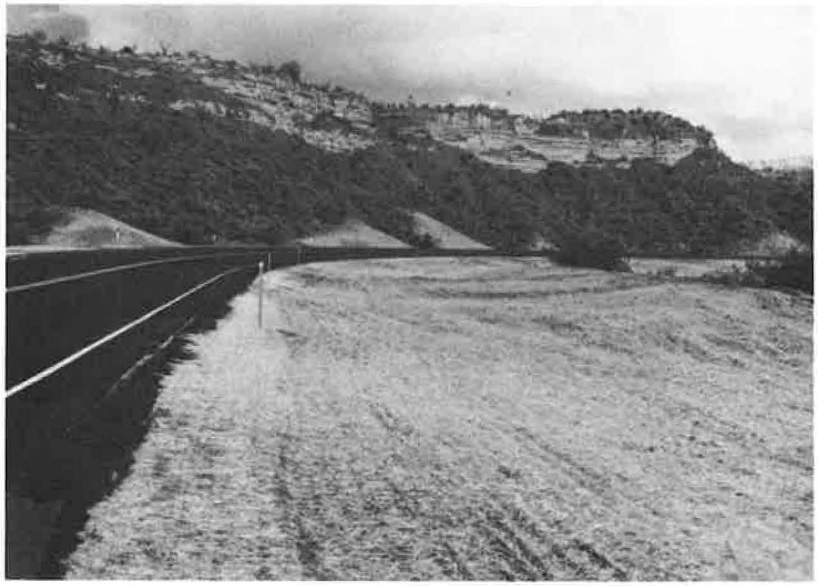

FIGURE 1 Rural roadside hazard rating of 1.

FIGURE 2 Rural roadside hazard rating of 3.

FIGURE 3 Rural roadside hazard rating of 5.

FIGURE 4 Rural roadside hazard rating of 7.