TRANSPORTATION RESEARCH

RECORD

No. 1401

Safety and Human Performance

Highway and Traffic Safety and Accident Research, Management, and Issues

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1993

Transportation Research Record 1401
ISSN 0361-1981
ISBN 0-309-05474-5
Price: $28.00

Subscriber Category
IVB safety and human performance

Printed in the United States of America
National Research Council. Transportation Research Board.

Sponsorship of Transportation Research Record 1401

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
Chairman: Jerome W. Hall, University of New Mexico

Facilities and Operations Section
Chairman: Jack L. Kay, JHK & Associates

Committee on Operational Effects of Geometrics
Chairman: Daniel S. Turner, University of Alabama
Secretary: Douglas W. Harwood, Midwest Research Institute
R. F. Benekohal, Myung-Soon Chang, John E. Fisher, J. L. Gattis, Don Jay Gripne, Michael F. Hankey, Raymond A. Krammes, Ken Lazar, Mark A. Marek, Patrick T. McCoy, John J. Nitzel, K. F. Phillips, Abishai Polus, Justin True, Walter E. Witt, Paul H. Wright

Committee on Methodology for Evaluating Highway Improvements
Chairman: Forrest M. Council, University of North Carolina
Secretary: Charles Philip Brinkman, U.S. Department of Transportation
Karen K. Ajluni, William T. Baker, William D. Berg, Brian L. Bowman, C.W. Colson, Jr., William D. Glauz, Alfred-Shalom Hakkert, Jerome W. Hall, Fred R. Hanscom, Ezra Hauer, David L. Helman, Julia L. Higle, Warren E. Hughes, Ruediger Lamm, Sidney J. Louick, King K. Mak, Thomas L. Maleck, Ernst Meyer, Olga Pendleton, Bhagwant N. Persaud, Jerry G. Pigman, Ramey O. Rogness, Harold T. Thompson, James M. Witkowski

Users and Vehicles Section
Chairman: Conrad L. Dudek, Texas A&M University

Committee on Operator Education and Regulation
Chairman: John W. Palmer, St. Cloud State University
Secretary: Charles B. Stoke, Virginia Department of Transportation
James E. Aaron, Edward W. Bleakly, Michael R. Calvin, Julie A. Clark, Eileen E. Eckhart-Strauch, Melvin Holmes, Donn W. Maryott, John M. McCullough, A. James McKnight, Kenard McPherson, Thomas L. Miller, Carol Lederhaus Popkin, Deborah A. Quackenbush, Chitaranjan Saran, Pete Shaffron, Michael F. Smith, E. Wallace Timmons, Michael A. Unger, Katie N. Womack, David M. Zaidel

Committee on Alcohol, Other Drugs, and Transportation
Chairman: Barry M. Sweedler, National Transportation Safety Board
Secretary: Carol Lederhaus Popkin, Natural Resources Environmental Health
William E. Collins, Grady C. Cothen, Jr., Frances Baker Dickman, Steven W. Gust, James H. Hedlund, Ralph Hingson, John H. Lacey, Hans Laurell, Roy E. Lucke, Adrian K. Lund, A. J. McLean, Herbert Moskowitz, James F. O'Hanlon, Barton Pakull, Raymond C. Peck, M. W. Bud Perrine, H. Laurence Ross, Monroe B. Snyder, Anthony C. Stein, Lawrence R. Sutton, Milan R. Valverius, Evelyn Vingilis, Robert B. Voas, Patricia F. Waller, Elisabeth N. Wells-Parker, R. Jean Wilson

Committee on Traffic Records and Accident Analysis
Chairman: Paul P. Jovanis, University of California-Davis
Secretary: Jeffrey F. Paniati, Federal Highway Administration
Sherry S. Borener, David J. Bozak, Thomas E. Bryer, Myung-Soon Chang, Benjamin V. Chatfield, Ted Chira-Chavala, Barbara Hilger DeLucia, Mark Lee Edwards, C. Arthur Geurts, Alan F. Hoskin, Patricia S. Hu, John J. Lawson, Joseph C. Marsh IV, Judson S. Matthias, Carol Lederhaus Popkin, Howard S. Stein, John G. Viner, Phyllis E. Young

Task Force on Statistical Methods in Transportation
Chairman: Olga Pendleton, Texas A&M University System
Secretary: Richard M. Weed, New Jersey Department of Transportation
Charles E. Antle, Lynda T. Carlson, Michael I. Darter, Gary A. Davis, Michael S. Griffith, Isit Hocherman, Patricia S. Hu, Robert L. Mason, Edward V. Milton, Frank C. Sonsini, Gary S. Spring, Peter R. Stopher, Robert E. White

GROUP 4—LEGAL RESOURCES
Chairman: Darrell W. Harp, New York Department of Transportation

Committee on Tort Liability and Risk Management
Chairman: Gordon S. Baca, Attorney
Chester J. Andres, Joseph D. Blaschke, Robert C. Blumenthal, Grady Click, C. Arthur Geurts, Gary L. Gittings, John C. Glennon, Brelend C. Gowan, Darrell W. Harp, David S. Johnson, Charles Raymond Lewis II, Russell M. Lewis, G. Rex Nichelson, Jr., Richard Arnold Olsen, Theresa A. Pasek, Norman L. Sisson, Jay L. Smith, Ruth W. Stidger, Kent Syverud, Michael E. Tardif, Rich Tiemeyer, Daniel S. Turner, Martin A. Wallen

Richard F. Pain, Richard A. Cunard, James B. McDaniel, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1992.

Transportation Research Record 1401

Contents

Foreword

"Traffic safety" is often mistaken to indicate a narrow technical field. As the papers in this Record illustrate, traffic safety encompasses a wide range of technical and management expertise, often technically and analytically complex. In this Record the papers range from discussions of safety management, safety evaluation, safety countermeasures, and accident analysis and prediction to tort liability. The papers further illustrate the current wisdom that there are no "silver bullets" in creating a safer highway transportation system. Many technical avenues must be explored and evaluated to most effectively bring about safer road and vehicle design and driver performance.

Assessment of Risk Management Procedures and Objectives in State Departments of Transportation

MICHAEL J. DEMETSKY AND KATHY YU

State departments of transportation (DOTs) throughout the nation face an increasing problem of compensation for damages caused by inaction or careless or negligent actions by their employees. When a state must compensate for such damages, it is said to be liable in tort. Many states have developed operational procedures to minimize loss due to tort liability; the procedures are generally referred to as risk management programs. The responses to a nationwide survey on risk management for tort liability by state DOTs are presented. The survey was designed first to determine the status of tort liability (i.e., which actions of the state and its employees constitute liability) and secondly the status of the risk management program (i.e., whether the state has established a formal risk management program to avert liability for its actions and those of its employees). If a state had established a tort liability risk management program, the details of the program were investigated to identify the key tasks of the program, including determination of objectives and criteria for measuring the effectiveness of a program, identification of hazardous situations, action taken upon notification of a hazardous situation, prioritization for mitigating hazardous situations, and documentation of notices, situations, and actions taken. A profile of a typical risk management program is provided. The profile can be used as a basis for developing a formal methodology by which DOTs can begin to evaluate existing or proposed risk management programs.

State departments of transportation (DOTs) throughout the nation face an increasing problem of compensation for damages caused by inaction or careless or negligent actions by their employees. When a state must compensate for such damages, it is said to be liable in tort. Until 1946, the doctrine of sovereign immunity protected states and their agencies from liability for death, injury, and property damage resulting from the negligent design, construction, or maintenance of roadways. With the passage of the Federal Tort Claims Act of 1946 and the corresponding erosion of sovereign immunity in many states either by statute or judicial decision, states began to find themselves open to tort liability.

Many states have developed operational procedures to minimize loss due to tort liability; the procedures are generally referred to as risk management programs. Although the number of elements in typical programs may vary, their general purposes are similar. The basic components of a risk management program include identifying risks, measuring the risks (evaluation), determining the method of controlling risks, implementing the method, and monitoring the results.

University of Virginia, Thornton Hall, Charlottesville, Va. 22903-2442.

The purpose of this paper is to describe the status of risk management programs in state DOTs and to indicate appropriate goals, objectives, and criteria for guiding their development. The information provided is intended to aid states that are in the process of establishing or evaluating their organizational structure for risk management. A survey of risk management practices in state DOTs provided most of the data for this study.

RISK MANAGEMENT SURVEY

In November 1989 a survey of risk management for tort liability was sent to 50 state DOTs and the Department of Transportation of the District of Columbia. Thirty-seven states and the District of Columbia responded to the survey. For the purpose of closure of the study, it was assumed that the no-response states had not established a formal risk management program at the time of the survey or that the survey did not reach an appropriate person. The survey was distributed to CEOs of state DOTs unless an individual responsible for risk management was known to the researchers. The initial distribution resulted in 31 responses. This was followed by a second mailing to the nonrespondents and a follow-up phone call. The final number of questionnaires used in this evaluation was 38.

The following sections describe the answers given by the respondents and provide a picture of how risk management for tort liability is presently addressed by state DOTs throughout the nation. The analysis of the survey is designed to assist the states in establishing or evaluating their organizational structure for risk management.

The questionnaire used for the risk management survey is available in the survey summary report (*1*). The questionnaire was designed to determine first the status of tort liability (i.e., which actions of the state and its employees constitute liability) and secondly the status of risk management programs (i.e., whether the state had established a formal risk management program to avert liability for its actions and those of its employees).

If a state had established a tort liability risk management program, the details of the program were investigated to identify the key tasks, including determination of objectives and criteria for measuring the effectiveness of a program, identification of hazardous situations, action taken upon notification of a hazardous situation, prioritization for mitigating

hazardous situations, and documentation of notices, situations, and actions taken.

STATUS OF TORT LIABILITY

It was found that all of the responding states except one had lost full sovereign immunity. To control the influx of claims and lawsuits, states have tried to protect themselves through statutory or judicial means. Some states have reinstated partial immunity such as state immunity, as a result of which the state cannot be held liable for its actions but its employees would be liable for their actions. Some states have employee immunity but not state immunity. Others have design immunity, as a result of which they can only be held liable for construction- or maintenance-related activities.

Another method of controlling tort liability costs has been to establish monetary limits for tort claims. For all the states responding, the limit for an individual does not exceed $250,000, for all states except New Hampshire, the limit for aggregate claims from an accident does not exceed $1,000,000. The limit in New Hampshire for aggregate claims is $2,000,000.

Judicially, the type of negligence law also influences tort liability cost. Negligence laws are either comparative or contributory (see Table 1). Most states indicated that they have a comparative negligence law. The states' answers as to whether having comparative law has helped or hurt the reduction of tort liability varied. The majority believed that having comparative law has increased tort liability. With contributory law, if the plaintiff is at fault, the whole claim is rejected without having to go to court. With comparative law, most of the claims reach court, and then the jury decides the case. In the end, states pay some amount for those cases.

STATUS OF RISK MANAGEMENT PROGRAMS

Of the 38 responding states (including the District of Columbia), 21 have a risk management program, and 3 are in the process of developing a risk management program. Still, all the states with risk management programs, with the exception of Missouri and Alabama, replied that they did not yet have a procedural manual for incorporating tort liability considerations in the design and maintenance processes. What they have implemented, generally, is a risk or safety management office or engineer to handle the problem. Various responses to the question, "Who is responsible for risk management in design and maintenance?" are given in Table 2.

RISK MANAGEMENT PROGRAM TASKS

The ensuing discussion focuses on tasks that are common to the evolving risk management programs. The following tasks were found to be similar in their mission, but, in many cases, different in form in the reported programs: hazardous situation identification, reaction to notification, prioritization, documentation, time limits (response), claims handling, criteria for litigation, and information systems.

Hazardous Situation Identification

Of 33 responses to this question, 31 states identify hazardous situations in three ways: citizen complaints, accident reports, and routine inspections. Twelve states also use other methods to identify hazardous situations. For example, one uses requests from attorneys and insurance companies, investigations for traffic control plans, and specific incidents as sources of information. Others use observations by their employees or police during routine activities or travel and information from legislators. Another state identifies hazardous situations through litigation trends, traffic engineering experience, and site investigations.

Reaction to Notification of a Hazardous Situation

Most of the states responded that a notice is either forwarded to the appropriate division for action or is corrected as soon as possible. Three states have procedures that are noteworthy here.

In one case, when notified of a hazardous situation, an accident summary is assembled. Then this summary is analyzed, and a collision diagram is prepared. On the basis of the information obtained, short- and long-term solutions are developed.

In another instance, as soon as notification is received, a "potential" file is created. All related documents are incorporated into this file. An assessment of what must be done to reduce liability is made, and then the names of any employees involved are collected. Also, photographs taken of the scene are added to the file.

TABLE 1 Definitions of Negligence

Contributory Negligence:	The plaintiff is barred from recovering damages for the accident for which he/she also was at fault.
Comparative Negligence:	The driver is not barred from collecting damages, because he/she was also at fault.
Pure Comparative Negligence:	A DOT could be required to pay the full amount of damages even if the plaintiff was 99% at fault and it was only 1% at fault.
Modified Comparative Negligence:	A plaintiff must prove that the DOT is over 50% at fault in order to recover any damages from the state DOT.

TABLE 2 Responsible Persons for Risk Management

State	Person
Alaska	Director of Risk Management
Arizona	Office of Risk Management
Colorado	Division of Risk Management
Hawaii	Assistant Chief of Construction and Maintenance
Idaho	Maintenance Supervisor, Traffic Supervisor, as well as Safety Program Coordinator
Iowa	Safety Review Engineer as well as Litigation Engineer
Louisiana	Department of Transportation and Development
Michigan	Supervisor of Litigation Coordination and Risk Management Section as well as Risk Management Engineer
Minnesota	Tort Claims Engineer
Missouri	Risk Manager
Oklahoma	Division Manager of Operations Review and Evaluation Division
Pennsylvania	Risk Management Engineer
Washington	Office of Risk Management
Wisconsin	Risk Manager

A third state performs an initial review, and the situation is then categorized either as a complex or simple matter. If simple, the following steps are taken:

1. The litigation coordinator/risk manager (LC/RM) coordinates and determines an action plan.
2. A memo is sent to the appropriate person ordering an immediate implementation of the action plan.
3. The LC/RM monitors the status of the action plan.

If complex, steps are more detailed:

1. The LC/RM coordinates risk evaluation.
2. A memo is sent to the appropriate persons asking for more detailed study and an action plan.
3. The LC/RM reviews the action plan devised by the responsible person to determine whether the plan achieves the risk management goals while remaining cost-effective.
4. If approved, the LC/RM sends the plan to upper management for funding approval. If approved, the action plan is implemented, and the LC/RM monitors the status of the plan.
5. If the LC/RM rejects the proposed plan, the plan must be revised. If the funding request is rejected by upper management, the plan must be revised.

Priority Determination

Of 25 states (including the District of Columbia), 3 reported that they use mathematical formulas to determine priorities:

- Iowa uses a composite rating based on accident rate, number of accidents, and the dollar loss.
- Colorado uses a weighted hazard index (WHI): $WHI = R_w - R_{wc}$, where R_w is the weighted accident rate and R_{wc} is the weighted critical accident rate (see Table 3).
- Texas uses a benefit/cost ratio to determine the priority of a situation (see Table 3).

Most other states use degree of hazard to determine priority. For example, in one case, a situation is classified as having one of four priorities: urgent (represents immediate hazard to the public and actions should be taken immediately), some danger to the public (the corrective actions should be taken as soon as possible during normal working hours), slight danger to the public (this should have a higher priority than regular maintenance activities), and finally, not urgent (this should be incorporated into the routine maintenance activities). Although examples of types of situations that fall into these categories were not given, an idea can be obtained from classifications used by other states. States tended to give life-threatening situations the highest priority and property damage situations the lowest priority. For example, Illinois gives highest priority to malfunctioning traffic signals, down Stop signs, snow and ice removal, pavement blowups, holes in bridge decks, shoulders lower than 3 in., and down Curve signs and No Passing signs. Illinois gives lowest priority to shoulders with less than a 3-in. drop, minor potholes, and delineators. The responses to the question are given in Table 4.

Documentation of Notices and Actions

Of the 33 states that responded to the question, 27 keep documentation of both the notices and the actions taken. Documentation, then, can be used in defense of the state's actions in a tort liability case.

Time Limits for Responses

Of the 32 states responding to this question, 14 post a time limit for corrective actions for reported defects. The time limit is established on the basis of potential degree of hazard. For example, one responded that for a traffic signal malfunction or Stop sign down, actions should be taken within 24 hr, and on less serious defects the statutory notice is 30 days. Another responded that the time limit was based on not only the type of defect but also on the location of the defect. For example, high-priority defects such as a knocked-down Stop sign would require immediate response. However, if the location is in a low-traffic area, the priority is reduced and the response does not have to be as quick. In another case, debris and spill on highways, regulatory and warning signs down, and storm dam-

TABLE 3 Prioritization Methods

1. Colorado:

$$WHI = R_w - R_{wc}$$

where

$$R_w = \frac{A_w}{VMT} \quad (1)$$

where

$$VMT = \frac{[(ADT) \times (\text{Section Length}) \times (\text{\# days in time period})]}{10^6} \quad (2)$$

A_w = PDO + (5 x INJ) + (12 x FAT)

PDO = number of property damage only accidents

INJ = number of injury accidents

FAT = number of fatal accidents

$$R_{wc} = R_{wa} + 1.5 \times \left(\frac{R_{wa}}{VMT}\right)^{1/2} - \frac{1}{(2 \times VMT)} \quad (3)$$

where

R_{wa} = statewide weighted average accident rate for the highway class in question

2. Texas:

$$\text{Safety Improvement Index (SII)} = \frac{B}{C} \quad (4)$$

where

C = initial cost of the project

B = present worth of project benefits over its service life

where

$$B = \frac{(S + 0.5xQ)}{1.08} + \sum_{i=2}^{L}\left[\frac{(S + 0.5xQ) + (i-1)xQ}{(1.08)^i}\right] \quad (5)$$

$$S = \frac{R \times C_f \times F + C_i \times I + C_p \times P}{Y} - M \quad (6)$$

$$Q = \left(\frac{A_a - A_b}{A_b} + L\right) \times S \quad (7)$$

where

S = annual savings in accident costs

R = percentage reduction factor

F = number of fatalities

C_f = cost of a fatality

I = number of injuries

C_i = cost of an injury

P = number of property damage only (PDO) accidents

Cp = cost of a PDO

Y = number of years of accident data

M = change in annual maintenance costs for the proposed project relative to the existing situation

Q = annual change in accident cost savings

A_a = projected average annual ADT at the end of the project service life

A_b = average annual ADT during the year before the project is implemented

L = project service life

age are corrected as soon as possible. Potholes are patched within 48 hr, and traffic control signs are corrected immediately. In a further instance, a time limit is established on the basis of potential hazard, corrective actions required, and availability of personnel.

Claims Handling

Of the 36 states responding, 32 keep records of all claims. Twenty of the 36 states classify claims for further use. Of the 16 states out of 20 that specified how the data were being used, 14 use the information to establish risk management priorities, and 12 use the information as input to decision making for functional activities such as routine maintenance, safety programming, and traffic engineering. Seven states responded that the information is used to implement actions taken at the statewide level, and three responded that information is sometimes used to implement actions for specific sites.

Criteria for Litigation

States will normally settle if the probability of losing the case is high. Some states will litigate even if the probability of losing is high due to the unreasonable amount of settlement demanded by the plaintiff. Some of the other factors taken into consideration in the determination of whether to litigate or to settle are persuasiveness of witnesses, ability of plaintiff's attorney, reputation of the judge assigned, issues of law and precedents, potential monetary loss, cost of litigation, potential for settling precedence, and public perception.

Information Systems

Twenty-two states responded that they process accident information to identify hazardous situations. Accident reports are used to determine areas that must be investigated and to summarize hazardous elements. For example, in some states, accidents are recorded by district, region, type, and cause. This permits identification of trends and potential deficiencies warranting special review or investigation. Also, in 21 states, once a potential trend or deficiency is identified, jurisdictions throughout the state are immediately notified.

Of 26 states reporting, 18 store information on centralized information systems, 15 store the information in centralized accident files, and 15 store information on local information systems. Fourteen states use more than one form of data storage for risk management.

Additional Strategies

In California, presentations on tort liability are given to any interested groups, and that state is beginning to incorporate risk management concepts in management performance evaluations. In Texas, a short course is being taught to employees regarding risk management to reduce roadway tort liability. Other states commented on the importance of educating em-

TABLE 4 Priority Determination

State	Method
Arizona	Degree of exposure; severity over frequency
Arkansas	Degree of hazard
Colorado	Weighted hazard index $WHI = R_w - R_{wc}$ where R_w is weighted accident rate and R_{wc} is weighted critical accident rate. See Table 3 for further detail
Dist. of Columbia	All equal priority unless life-threatening
Hawaii	Based on safety, health and welfare of public
Idaho	Highest: life-threatening Lowest: problems not directly in traffic areas
Illinois	Highest: malfunctioning traffic signals, down stop signs, pavement blowups, holes in bridge deck, shoulders lower than three inches, down curve signs and no passing signs. Lowest: shoulders with less than a three inch drop, minor pot holes and delineators
Iowa	Composite reading based on accident rate, number of accidents, and dollar loss
Kentucky	Benefit/Cost ratio
Michigan	Based on safety and payouts
Minnesota	Degree of Hazard
Missouri	Priority 1: Urgent. Represents immediate hazard to public. Should respond as soon as possible. Priority 2: Some danger to the public. Should be accomplished as soon as possible during normal hours Priority 3: Slight danger to public. Repair should be accomplished with higher urgency than routine maintenance Priority 4: Not urgent. Considered common occurrence with no danger to public. Would normally be considered routine maintenance.
New Hampshire	Degree of hazard
Ohio	Seriousness of hazard
Oklahoma	Degree of hazard
Pennsylvania	Degree of hazard, exposure to risk, competing needs, availability of manpower, available funds, etc.
Rhode Island	The degree of crisis dictates priority. All situations are addressed within 48 hours.
South Dakota	First come, first served basis
Tennessee	Highest: life-threatening Lowest: property damage
Texas	Mathematical formula
Virginia	Judgment call by field engineers
Washington	Highest priority: malfunctioning traffic control devices and damaged road surface, and snow and ice removal
West Virginia	Case by case
Wisconsin	Prioritized weekly
Wyoming	Severity of injury

ployees regarding tort liability so that they will be conscious of it while they are performing their duties. The additional comments provided are summarized in Table 5.

RISK MANAGEMENT PROGRAM OBJECTIVES

The organizational structure and elements of a state DOT's risk management program are a reflection of the particular objectives for risk management that have been established in that jurisdiction. The following is a summary of the survey responses regarding risk management objectives:

- To improve highway safety by identifying, analyzing, prioritizing, and recommending alternatives to change the roadway environment in a manner that will reduce motor vehicle accidents;
- To reduce the department's exposure and loss due to liability;
- To coordinate and track all claims and litigation against the department, to process claims and manage a tort liability loss-mitigation program, and to direct the resources of the department to minimize the adverse effects of litigation on the department and the public;
- To serve as the tort claim representative for the department and coordinate investigative service with the attorney general's office; and
- To administer an employee safety program, to promote a cost-effective risk management effort statewide, to develop control mechanisms through training and counseling, and to

TABLE 5 Concepts Not Directly Addressed in Survey

State	Comments
California	Presentations made to a variety of interested groups on the subject of tort liability. Experiences and trends in the law are given at various meetings, conferences and training sessions for traffic, design, maintenance and construction employees.
Colorado	Colorado has "mandatory arbitration" for certain types of cases.
Kentucky	Currently conducting a research project to review tort claims against the Kentucky Transportation Cabinet and provide information to use in establishing a risk management program.
Louisiana	The Office of Risk Management acts as carrier for all state agencies including Department of Transportation and Development (DOTD). Its claims section handles all tort actions against state, and its loss prevention section oversees various safety and loss prevention activities and programs. The DOTD does have internal procedures for prioritizing and acting to correct potentially hazardous situations.
Minnesota	Employee training to explain the litigation procedure, as well as the importance of following specific design, maintenance, and construction policies and procedures, and the responsibility of proper documentation of actions taken in the field.
Missouri	Employee safety, hazardous material management, and property damage to state property.
South Dakota	Numerous training programs, defensive driving, safety evaluations, and certifications.
Texas	A short course titled "Risk Management to Reduce Roadway Tort Liability" is taught periodically throughout the state by Texas Transportation Institute, Texas A&M University.
Wisconsin	Workmen's Compensation, Hazardous Material Management, Safety Management, Fleet Liability

foster an awareness by all employees of the risk potential associated with their actions.

As indicated, there is a great variety in the goals and objectives as determined from the survey. Some are more clearly defined than others, some are more generic, and some are very focused.

RISK MANAGEMENT PROGRAM CRITERIA

Of the 24 states responding, 23 use more than one criterion to measure the success of their program. Twenty-two states use the total number of claims as one of the criteria, 21 use the cost of all claims paid, 19 use number of claims paid, and 14 use number of accidents as one of the criteria. Only two states responded that their programs are not evaluated. Five states responded that they use criteria other than those stated above. In the survey, one state also uses cost per claim, number of improvements completed, and standards or policies revised. Another uses safety improvement potential of proposed changes in policy and procedures as well as the other criteria, and one uses the reactions and opinions of DOT defendants. Overall, these criteria ranged from highly objective and measurable to quite subjective.

FRAMEWORK FOR RISK MANAGEMENT PROGRAMS

The results of the risk management survey that have been described constitute a basis for the development of a methodology to evaluate proposed and existing risk management programs. They can also be used to assist in defining the organizational and functional needs for a state DOT's risk management program. These risk management elements and tasks are interrelated as indicated in Table 6. Table 6 also gives criteria that are appropriate for measuring the effectiveness of each element in terms of aggregate program performance rather than task performance.

Profile of a Typical Risk Management Program

Using these building blocks, a typical approach to risk management is derived from the state of the practice.

The states typically identify hazardous situations by citizen complaints, accident reports, and routine inspections. Once notified of a hazardous condition, a state usually has an established procedure for action. Only three states use a mathematical model for prioritization of defects for remedial action. Most states use a subjective degree of hazard to determine priorities.

After receiving a notice from a citizen concerning a possible defect, 17 of 33 states follow up any action with a call to the informant. Most states maintain documentation of the notices received and actions taken. Nearly half of the respondents have established time limits for taking action to mitigate different types of reported defects or problems. Time limits are based on the potential degree of hazard, ranging from 24 hr for traffic signals to 30 days for less serious problems.

The majority of states keep records of all claims; however, many do not classify them according to type of hazard to provide direct information for needed areas of improvement. Some only classify settled claims. In most cases, this information is not used to evaluate and establish risk management

TABLE 6 Building Blocks for Risk Management

Elements	Tasks	Criteria
Risk Identification	• Hazardous system identification	• Number of hazardous situations identified; by employees; through citizen complaints
Risk Evaluation	• Prioritization • Ranking	• Seriousness of injuries • Number of accidents • Number of claims
Risk Control	• Reaction to notification • Follow-up • Time limits	• Number of claims • Number of situations corrected • Number of accidents • Number of fatalities
Implementation	• Objectives • Organizational structure • Governing legislation • Responsibility • Policy & procedures • Support systems • Claims handling • Documentation • Transfer	• Response time
Monitoring & Feedback	• Criteria • Information systems	• Total cost • Cost of resources • Cost of tort liability

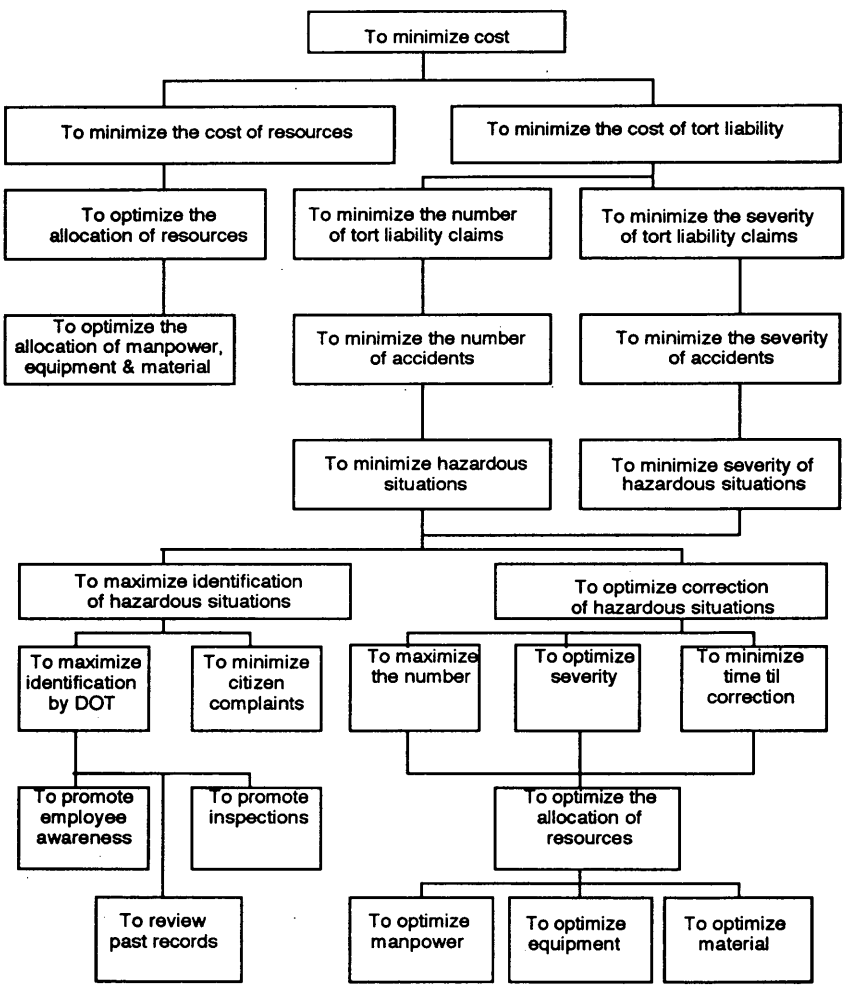


FIGURE 1 Goals and objectives of a risk management program.

policy. Normally, subdivisions throughout a state are notified of a particular defective situation.

Various criteria are used by states to decide whether to settle or litigate a case. Basically, if it appears that a case will be lost, it is settled. However, if a claim is unreasonably high, it may be litigated even though the chances of winning are slim.

Most states process accident incident information to identify the existence of hazardous situations. Data are typically available on all or combinations of the following: centralized information system, central accident files, and local information systems. At the time of the survey, many states were training employees with risk management procedures to reduce tort liability. Awareness is the objective.

Development of Goals and Objectives for Risk Management

At this point, it is appropriate to withdraw from the constructive details of the risk management program and to define a scope of the goals and objectives that are needed to begin to structure a program. What should a program accomplish? Since it is designed to minimize tort liability costs, that should be one of its main objectives. However, states must also realize that in minimizing tort liability costs, they do not have unlimited resources. This requires that they balance minimizing the cost of tort liability with the cost of resources. Figure 1 shows a possible hierarchial goal structure of a risk management program that takes these considerations into account. This effort is enhanced by the rapid identification of hazardous situations and their mitigation through manpower, equipment, and materials. Concepts such as those shown in Figure 1 should be investigated and associated analytical processes developed that provide a method for structuring and evaluating risk management programs in state DOTs. These methodologies would build on the ideas presented in this paper in consideration of program goals, objectives, and criteria that identify and evaluate the primary tasks that constitute the program.

CONCLUSIONS

The preceding has been a summary of the common practices regarding risk management for tort liability in state DOTs. These observations provide a profile of risk management practice. When risk management programs become more clearly defined, the use of available information to develop risk mitigation strategies will dominate program improvements. The critical information areas are (*a*) identification of defective situations; (*b*) association of these defects with design, construction, operational, and maintenance practices; (*c*) communication of problems and solutions throughout the agency; and (*d*) maintenance of a comprehensive data base where claims are classified in such a way that sites, hazards, and remedial actions are identifiable. These findings can be used to structure a formal methodology for evaluating existing or proposed risk management programs.

ACKNOWLEDGMENTS

This research was sponsored by the U.S. Department of Transportation, University Transportation Centers Program, Region III, and the Virginia Department of Transportation. The authors are indebted to the states that provided the information for the study through the survey.

REFERENCE

1. K. Yu and M. J. Demetsky. *Risk Management Systems, Volume I: A Survey of Risk Management in State Departments of Transportation.* Report UVA/529685/CE91/104. Department of Civil Engineering, University of Virginia, Charlottesville, Va., May 1991.

Publication of this paper sponsored by Committee on Tort Liability and Risk Management.

Evaluation of Safety Impact of Highway Projects

HASHEM R. AL-MASAEID, KUMARES C. SINHA, AND THOMAS KUCZEK

An empirical Bayesian approach to evaluate the safety impact of highway projects at a group of sites level was developed. Rural traffic accident data from the state of Indiana were used. The Bayesian methodology was illustrated using examples of wedge and level and resurfacing projects. The results indicated that wedge and level and resurfacing activities did not have a significant effect on the level of the expected number of accidents or accident rates at 95 percent probability level on Indiana's two-lane rural roads having traffic volumes less than 4,000 vehicles per day.

The purpose of the research reported in this paper is to develop an empirical Bayesian approach to evaluate the safety impact of highway projects in reducing accident frequency, accident severity, or the potential for accidents. The results are used to determine accident reduction factors (ARFs).

Four methods were reviewed. First, the simple before-and-after method is based on the observed number of accidents. Because of the regression-to-mean phenomenon, this method tends to overestimate the effectiveness of highway safety improvements (*1*). Second, the before-and-after-with-matched-control-group method, although theoretically sound, is difficult to apply in practice (*2*). The third method, a modeling technique, cannot completely explain all variations in traffic accidents; thus its use is limited (*3*). And fourth, the Bayesian approach was only partially addressed by Hauer (*4,5*) because his evaluation was based on comparisons observed in the after period with the predicted values at site level. Moreover, probability levels were not computed to declare any significance in the results.

The present study applied Bayesian statistics to estimate the expected number of accidents (mean in the long term) and the expected accident rate. Bayesian statistics were used for three main reasons. First, the Bayesian approach is a probabilistic method capable of augmenting the most recent information with the available historical data or prior knowledge to achieve better estimates, reducing associated uncertainties. Second, the Bayesian approach permits pooling of information for a population or set of sites as well as for individual sites, allowing better use of the available information for prediction purposes. Third, the performances of the Bayesian models to predict the expected number of accidents and accident rate were investigated (*6*), and they provided legitimate estimates specifically at a group of sites level.

H. R. Al-Masaeid, Jordan University of Science and Technology, Irbid, Jordan. K. C. Sinha and T. Kuczek, Purdue University, West Lafayette, Ind.

MODEL DEVELOPMENT

In the analysis, two assumptions were made:

1. Traffic accidents at any particular location fit the Poisson distribution in the absence of any highway improvement.
2. The expected number of accidents λ is a random variable with a gamma probability distribution over the population of a number of sites. The expected accident rate ρ is a random variable with a gamma probability distribution.

On the basis of the assumptions, if x denotes the observed number of accidents at a particular location in 1 year, x has a Poisson distribution $P(x/\lambda)$ with mean λ so that

$$P(x/\lambda) = \lambda^x e^{-\lambda}/x! \qquad x = 0, 1, 2 \tag{1}$$

Considering λ (mean in the long run) as a continuous random variable within the population of locations and $f(\lambda)$ as its probability density function with parameters β and α, the prior distribution of λ is

$$f(\lambda) = a^\beta \lambda^{\beta-1} e^{-\alpha\lambda}/\Gamma(\beta) \qquad \lambda, \beta, \alpha > 0 \tag{2}$$

where $\Gamma(\beta)$ is the gamma function.

The parameters of the prior distribution are estimated from the sample statistics of similar locations as follows:

$$\hat{\alpha} = \bar{x}/(s^2 - \bar{x}) \tag{3}$$

$$\hat{\beta} = \bar{x}^2/(s^2 - \bar{x}) \tag{4}$$

where $\bar{x}$ and s^2 are the sample mean and sample variance calculated from all similar sites in the population.

The next step in the analysis is to combine the prior knowledge and the new information to obtain the posterior distribution of λ. This is accomplished through the application of the Bayes' theorem. Thus, if x is the number of accidents at a given location, the posterior distribution of λ is of gamma type with parameters

$$\beta' = x + \beta \tag{5}$$

and

$$\alpha' = 1 + \alpha \tag{6}$$

Having obtained the posterior distribution of λ at each location, the next step is to obtain the distribution of the total

expected number of accidents λ_t at a group of n similar locations. This is accomplished by using the convolution principle. Thus, if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are independent random variables, the total expected number of accidents λ_t is given by

$$\lambda_t = \sum_{i=1}^{n} \lambda_i \tag{7}$$

where λ_i is the expected number of accidents at location i.

Moreover, convolutions of n probability density functions each having posterior parameters of β_i' and α' have a gamma probability density function (*7*). Therefore, λ_t has a gamma probability density with parameters $\Sigma\beta_i'$ and α', as shown below:

$$\sum \beta_i' = n\beta + \sum x_i \tag{8}$$

$$\alpha' = 1 + \alpha \tag{9}$$

The expected mean and variance of λ_t are

$$E(\lambda_t) = \left(n\beta + \sum_{i=1}^{n} x_i\right)/(1 + \alpha) \tag{10}$$

$$Var(\lambda_t) = \left(n\beta + \sum_{i=1}^{n} x_i\right)/(1 + \alpha)^2 \tag{11}$$

Safety impacts may be estimated by comparing the number of accidents that probably would have occurred (according to the posterior distribution before improvement) with the actual observed at the location after the improvement. However, when doing so, the predictive distribution of the number of accidents will be of the negative binomial form (*8*), which has high variance. Therefore, a large reduction in the observed number of accidents in the after period is necessary to dictate any significant difference specifically at the site level.

The evaluation method can be enhanced by using the expected number of accidents, which has less variability than the number of accidents. To do so, it is necessary to assume that the expected numbers of accidents before and after improvements are independent random variables.

Thus, if β' and α' are the posterior parameters of the expected number of accidents in the before period λ_b, and the expected number of accidents in the after period λ_a has posterior parameters k' and γ', the joint probability density function is

$$f(\lambda_b, \lambda_a) = \begin{cases} \dfrac{\alpha^{\beta'}\lambda_b^{\beta'-1}e^{-\alpha'\lambda_b}}{\Gamma(\beta')} * \dfrac{\gamma^{k'}\lambda_a^{k'-1}e^{-\gamma'\lambda_a}}{\Gamma(k')} & \lambda_b, \lambda_a, \alpha', \beta', \gamma', k' > 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{12}$$

Therefore, to evaluate the safety effectiveness of a given improvement, one must compute the probability of the expected number of accidents in the after period, λ_a, being less than the expected number of accidents in the before period. That is,

$$p(\lambda_a < \lambda_b) > \theta \tag{13}$$

where θ is the predetermined probability level at which the improvement is declared effective in the reduction of the expected number of accidents.

The preceding probability statement can be expressed as follows:

$$p(\lambda_a < \lambda_b) = 1 - p(\lambda_a > \lambda_b)$$
$$= 1 - \int_{RS}\int f(\lambda_b, \lambda_a)d\lambda_b d\lambda_a \tag{14}$$

where RS is the region satisfying $\lambda_a > \lambda_b$. Since λ_a and λ_b take only positive values,

$$p(\lambda_a < \lambda_b)$$
$$= 1 - \int_0^\infty \int_{\lambda_b}^\infty \frac{\gamma^{k'}\lambda_a^{k'-1}e^{-\gamma'\lambda_a}}{\Gamma(k')} * \frac{\alpha^{\beta'}\lambda_b^{\beta'-1}e^{-\alpha'\lambda_b}}{\Gamma(\beta')} d\lambda_a d\lambda_b \tag{15}$$

The preceding integral is equal to

$$p(\lambda_a < \lambda_b)$$
$$= 1 - \sum_{j=0}^{k'-1} \left[\frac{\gamma'}{\alpha'}\right]^j \left[\frac{\alpha'}{\alpha' + \gamma'}\right]^{\beta'+j} * \frac{\Gamma(\beta' + j)}{\Gamma(j + 1)\Gamma(\beta')} \tag{16}$$

On the other hand, the use of the accident rate method to assess the safety impact of highway improvement is valuable because the accident rate method is sensitive to variations in traffic volume. The probability of x accidents at a site i with accident rate ρ and volume V_i is given by

$$P(x|\rho, V_i) = \frac{(\rho V_i)^x e^{-\rho V_i}}{x!} \tag{17}$$

where V_i is the normalized average daily traffic volume (ADT$*365/10^6$) at a given site.

The second assumption implies that the probability distribution function of the accident rate at a site has a gamma distribution; that is, the parameters of the gamma distribution can be estimated by matching the mean and variance of the observed accident rates [R and Var (R)] from a sample with the mean and variance of negative binomial distribution normalized by the volume. Morris (*9*) pointed out that the estimates of gamma parameters based on the method of moments are as follows:

$$\hat{v} = V^* * \overline{R}/(V^* * S^2 - \overline{R}) \tag{18}$$

$$\hat{a} = \overline{R}v \tag{19}$$

where

a and v = shape and scale parameters of gamma distribution, respectively;
V^* = harmonic mean of traffic volumes (normalized $V_1, V_2, \ldots, V_n$);
S^2 = variance of the observed accident rates obtained from a sample of sites; and
$\overline{R}$ = mean of accident rates in the population of sites ($R_1, R_2, \ldots, R_n$).

The mean and variance are computed from a sample as follows:

$$\overline{R} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} R_i \tag{20}$$

$$S^2 = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} (R_i - \overline{R})^2 \tag{21}$$

where R_i is the accident rate at Site i and n is the number of sites in the selected sample.

Once the parameters of prior distribution have been determined, the next step is to combine the prior information with the site-specific data to obtain the posterior distribution. Thus, if x_i is the number of accidents and V_i is the normalized traffic volume on a given site, the posterior distribution of ρ is of gamma form with parameters

$$\nu_i' = \nu + V_i \tag{22}$$

and

$$a_i' = a + x_i \tag{23}$$

At the group of sites level, the total expected accident rate is given by the sum of individual accident rates. Thus, the total expected accident rate ρ_t for a group of n sites is given by

$$\rho_t = \sum_{i=1}^{n} \rho_i \tag{24}$$

where ρ_i is the expected accident rate of Site i.

Moreover, the expected value and variance of ρ_t are

$$E(\rho_t) = \sum_{i=1}^{n} (a_i'/\nu_i') \tag{25}$$

and

$$Var(\rho_t) = \sum_{i=1}^{n} (a_i'/\nu_i'^2) \tag{26}$$

Parameters of ρ_t can be computed by matching moments from Equations 25 and 26. The parameters can be approximated as follows:

$$\nu = \left(\sum a_i'/\nu_i'\right) \Big/ \left(\sum a_i'/\nu_i'^2\right) \tag{27}$$

$$a = \left(\sum a_i'/\nu_i'\right)^2 \Big/ \left(\sum a_i'/\nu_i'^2\right) \tag{28}$$

The general approach used in the evaluation of safety impacts according to the expected number of accidents will be used herein. To evaluate safety impacts according to the accident rate model, it is assumed that accident rates before and after improvements are independent random variables. Hence, if ρ_b is the before accident rate with posterior parameters a' and ν', and ρ_a is the after accident rate with posterior parameters b' and u', then a highway project is efficient in reducing accident rate if and only if the probability that ρ_a is less than ρ_b exceeds a predetermined probability level (say θ). That is,

$$p(\rho_a < \rho_b) > \theta \tag{29}$$

where $p(\rho_a < \rho_b)$ can be approximated from the following derived equation:

$$p(\rho_a < \rho_b) = 1 - \sum_{j=1}^{b'-1} \left[\frac{u'}{\nu'}\right]^j \left[\frac{\nu'}{\nu' + u'}\right]^{a'+j} * \frac{\Gamma(a'+j)}{\Gamma(j+1)\Gamma(a')} \tag{30}$$

METHODOLOGY

In this research, a simple before-and-after methodology based on the Bayesian approach is presented to evaluate safety impacts of highway projects. In the evaluation both the expected number of accidents and the expected accident rate at a group of sites (total expected number of accidents and total expected accident rate) were used, rather than the expected number of accidents or expected accident rate at a site level because at site level a very large change in accidents is needed to judge significant results. The before-and-after methodology based on the Bayesian approach can be summarized as follows:

1. In the before period, prior parameters estimated from previous knowledge are augmented with the most recent information to estimate either the expected number of accidents or the expected accident rate. The resulting estimated value represents the best estimate of the expected value in the future period in the absence of any highway improvements. In the research, 2 years of accident experiences at all sites represented the before period. The prior parameters were estimated from the first year, whereas the sample data (sample likelihood) was drawn from the second year. The second year provided data immediately preceding the implementation of an improvement project.
2. In the after period, the prior parameters were estimated from the data of the first year after improvement. The prior parameters were augmented with the available information from the second year in the after period to estimate the posterior parameters and to predict either the expected number of accidents or the accident rate. If a large change in the average daily traffic volume is noted, adjustment of parameters according to the change in volume is necessary if the expected number of accidents is used. The predicted value based on the posterior distribution in this stage then represents the best estimate of the total expected number of accidents or the total accident rate after the installation of a highway improvement.

In fact, prior and posterior information need not be separated in time if there are a large number of similar sites to estimate prior parameters. However, similarity is a subjective matter. Therefore, it was believed that prior parameters estimated from the sites under investigation (that will receive or received an improvement) would narrow down the similarity and provide more reliable results. In this case, separation is necessary, at least from a theoretical point of view, because it is impossible to estimate prior and posterior pa-

rameters from the same set of data. In addition, the separation provided a means to incorporate the accident data within 2 years in each period. The regression-to-mean effect is defined as the difference between the posterior mean and the past observed mean. In this methodology, comparisons were carried between posterior means; therefore, the regression-to-mean effect is not a considerable issue.

3. Comparisons of predicted values from the before and after periods were used to estimate the percentage change in the values of the total expected number of accidents or total accident rate. Moreover, if the evaluation was based on the total expected number of accidents, Equation 16 was used to compute the probability that the total expected number of accidents in the after period is less than the total expected number of accidents in the before period for the group of improved sites. Similarly, Equation 30 was used to compute the associated probability level if the evaluation was based on the total expected accident rate.

DATA DESCRIPTION

Rural traffic accident data from the state of Indiana were used in the study. Traffic accident data from police records for 1982 through 1989 were used to estimate the safety impact of wedge and level and resurfacing projects. The wedge and level and resurfacing projects were selected mainly because of the availability of a large number of sites with these activities.

In the study, a site is defined as a section of rural highway 1 mi long irrespective of its past accident history. Sites of intersections and interchanges were excluded. Annual average daily traffic volumes on each site were obtained from Indiana Department of Transportation volume statistics publications. Moreover, detailed maps were extensively used to define the boundaries of each site or a group of consecutive improved sites.

Evaluation of the Safety Impact of Wedge and Level Improvement

Basically, wedge and level activity is not a safety improvement. However, in this paper its safety impact was evaluated as an example to present the methodology. Most of the wedge and level projects in the sample were implemented on non-Interstate highways in Indiana from 1983 through 1986. Almost all affected sections had annual average daily traffic volumes of less than 4,000 vehicles per day. A sample of 190 sites (190 mi) was selected to estimate the safety impact of wedge and level activity.

To apply the developed methodology, accident experiences of sites for 2 years before and 2 years after improvement were required. As mentioned earlier, the first year in the before period was used to estimate the before prior parameters, and the accident experience of all sites in the second year (immediately before the year of implementation) was used to update the prior parameters and to compute the posterior parameters. In the same manner, in the after period, accident experience in the year immediately after the year of implementation was used to estimate the prior parameters, and accident experiences of all sites in the second year were combined with prior to estimate posterior parameters and to predict the expected number of accidents or the accident rate.

Estimation of Parameters and Other Issues

For the estimation of prior parameters two samples were selected randomly from the population of improved sites in the before and after periods. Each sample had 60 sites. These samples were used to estimate prior parameters of the expected number of accidents and accident rate and then to show that the marginal distribution of the observed number of accidents approximately followed the negative binomial distribution and, as a result, the expected number of accidents or accident rate had a gamma probability distribution. The goodness-of-fit results indicated that the marginal distribution of the observed number of accidents followed the negative binomial distribution. Therefore, the assumption of the gamma distribution adopted in the derivation of Equations 16 and 30 is not unreasonable.

To augment the prior information, the sample data (sample likelihood) from all sites (190 sites) in the before and after period were used. The sample data were combined with prior parameters to predict the expected number of accidents or accident rates in the before and after periods. However, before the final assessment of the wedge and level activity, it was instructive to check for the assumption made in the derivation of Equations 16 and 30. The assumption stipulated that the expected number of accidents and accident rates in the before and after period are independent random variables. For this purpose, the posterior parameters for each site were computed from prior parameters and the site specific data before and after improvements. The results from the 190 sites are shown in Figures 1 and 2 for the number of accidents and accident rate, respectively. The estimated correlation coefficient between the before and after values of the expected number of accidents was 0.00396, and the corresponding coefficient for the accident rate was −0.00059. The correlations were very small and not significant; therefore, the assumptions of independence are not unreasonable.

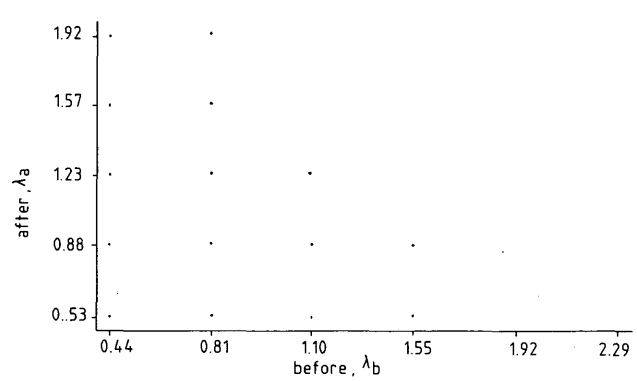


FIGURE 1 Expected number of accidents in the before and after improvements for wedge and level improvements.

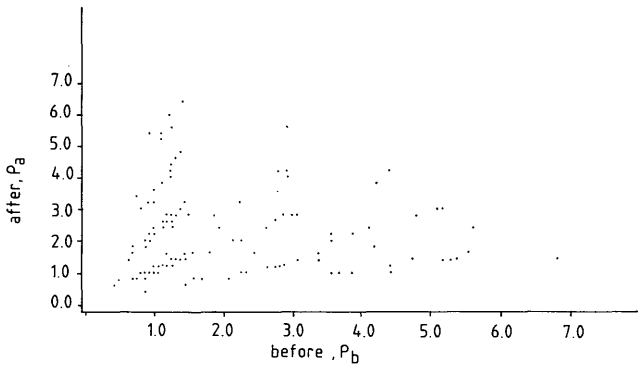


FIGURE 2 Expected accident rates in the before and after improvements for wedge and level improvements.

Results of Evaluation

Wedge and level activities were evaluated according to the expected number of accidents model and the accident rate model. The evaluations were performed at a group of sites level (population of the improved sites). Results of analyses according to the expected number of accidents model are given in Table 1, and results of evaluation according to the accident rate model are summarized in Table 2. Since there was a considerable increase in the total average daily traffic volume (the total volume on all sites changed from 181,475 vehicles per day in the before period to 188,890 in the after period), the decision with respect to the expected number of accidents was made on the adjusted results (transformation results). The adjustment was made as indicated in Table 1. The results indicated that the estimated mean number of traffic accidents increased by 8.06 after implementation of the wedge and level improvements. However, the probability test from Equation 16 indicated that the increase was not significant at 5 percent level [$P(\lambda_a < \lambda_b) = 0.131$]. The increase in the accident rate after the wedge and level improvement was 7.70 percent. The probability test according to Equation 30 indicated that the increase in the accident rate was not significant at 5 percent level [$P(\rho_a < \rho_b) = 0.188$].

Evaluation of the Safety Impact of Resurfacing Improvements

Resurfacing projects were implemented on both undivided and divided roads. However, only undivided U.S. and state roads had sufficient numbers of improved sites. The total number of sites on undivided U.S. and state roads that were used in the evaluation was 95 (95 mi). The annual average daily traffic volume on these sites was less than 4,000 vehicles per day in the year of implementation.

In the before period, prior parameters for the expected number of accidents and accident rate were estimated from accident experience and traffic volume in the first year of the 2 years preceding the resurfacing. In the after period, prior

TABLE 1 Estimation of Accident Reduction Factor Associated with Wedge and Level Improvements According to the Expected Number of Accidents Model

Prediction for Period		Before period	After Period
Prior Parameters	α	1.7073	1.8862
	ß	1.1950 (1.3737)*	1.5400
Sample Data	$\sum n_i$	190	190
	$\sum x_i$	115 (119.7)**	146
Convoluted Parameters	α'	2.7070	2.8862
	ß'	342.00 (380.71)*	439.00
Traffic Volume	Prior	66695	76670
	Sample	181475	188890
Predicted Values		126.36 (140.66)*	152.00
Accident Reduction Factor	Unadjusted=((126.36-152.00)/126.36)*100=-20.30 %		
	Adjusted=((140.36-152.00)/140.36)*100=-8.06%		

* Adjusted according to traffic volumes.
** 115 · (188890/181475).

TABLE 2 Estimation of Accident Reduction Factor Associated with Wedge and Level Improvements According to the Accident Rate Model

Prediction for Period		Before period	After Period
Prior Parameters	α	0.714325	0.863700
	ν	0.414030	0.493265
Sample Data	$\sum n_i$	115	146
	$\sum x_i$	66.2384	68.9450
Expected Accident rate	MEAN	354.7944	382.0959
	VARIANCE	541.1785	500.8576
Estimated Parameters	α'	333.00	291.00
	ν'	0.65650	0.7620
Accident Reduction Factor	A.R.F.=((354.7944-382.0959)/345.7944)*100=-7.70 %		

parameters for the expected number of accidents and accident rate were estimated from the first year following the improvement. In the before period the sample data were from the year preceding the implementation, and in the after period the sample data were from the second year after implementation. Posterior parameters were used to predict the expected number of accidents and rate of accidents in the before and after periods. On the basis of the posterior parameters, the correlation between the expected number of accidents and the expected accident rates in the before and after periods were 0.0943 and 0.11077, respectively. These correlations were not significant at 5 percent level.

Results of Evaluation

Resurfacing improvements were evaluated according to the expected number of accidents and accident rate models. Results of the evaluation based on the expected number of accidents model are summarized in Table 3. Results of the evaluation according to the accident rate model are summarized in Table 4.

Table 3 indicates that the level of traffic accidents increased after implementation of the resurfacing projects. However, the total average traffic volume also increased from before to after periods (from 212,195 to 235,872 vehicles per day). The accident number was therefore adjusted for traffic volumes. The adjusted results indicated the increase in the expected number of accidents to be about 7.66 percent. According to Equation 16 the increase was not significant at 95 percent level [$P(\lambda_a < \lambda_b) = 0.163$]. Table 4 indicates that the increase in the rate of accidents in the after period was 11.10 percent. However, the computed probability level for Equation 30 indicates that with 5 percent probability level the increase was not significant [$P(\rho_a < \rho_b) = 0.081$].

DISCUSSION OF RESULTS

This paper presented a methodology to evaluate the safety impact of highway projects based on the Bayesian approach. The methodology was applied to evaluate two types of highway projects on rural undivided U.S. and state roads having annual average daily traffic volumes of less than 4,000 vehicles per day. The examples were drawn from wedge and level projects and resurfacing projects. These projects are not necessarily safety improvement projects, and the results of analyses indicated that there was no significant difference in the level of traffic accidents before and after implementation of wedge and level and resurfacing improvements.

The aim of wedge and level improvements is to wedge rutting depression and holes in the pavement and to provide acceptable leveling of the roadway cross section. Longitudinal rutting depression on roadway sections causes water and snow or ice to accumulate on the pavement surface, and as a result vehicles are more likely to be involved either in hydroplaning accidents or in slippery pavement accidents. Rutting depression, transverse depression, and excessive lack of roadway leveling are main causes of loss of control accidents and driver fatigue accidents. Thus, the promise of wedge and level as safety improvements is to reduce these types of accidents and to provide better driving conditions. Results of analyses indicated that the number of accidents and the rate of accidents increased after implementation of these improvements. The

TABLE 3 Estimation of Accident Reduction Factor Associated with Resurfacing Improvements According to the Expected Number of Accidents Model

Prediction for Period		Before period	After Period
Prior Parameters	α	1.6683	1.9100
	ß	1.7934 (1.9323)*	2.2920
Sample Data	$\sum n_i$	95	95
	$\sum x_i$	104 (115.0)**	132
Convoluted Parameters	α'	2.6683	2.9100
	ß'	274.40 (298.00)*	350.00
Traffic Volume	Prior	82640	89045
	Sample	212195	235872
Predicted Values		102.83 (111.63)*	120.19
Accident Reduction Factor	Unadjusted=((102.83-120.19)/102.83)*100=-16.88 %		
	Adjusted=((111.63-120.19)/111.63)*100=-7.66%		

* Adjusted according to traffic volumes.
** 104 * (235,872/212,195)

TABLE 4 Estimation of Accident Reduction Factor Associated with Resurfacing Improvements According to the Accident Rate Model

Prediction for Period		Before period	After Period
Prior Parameters	α	5.286000	1.200000
	ν	3.710000	0.810300
Sample Data	$\sum n_i$	104	132
	$\sum x_i$	77.4512	86.0933
Expected Accident rate	MEAN	134.1503	149.0442
	VARIANCE	29.9145	97.1015
Estimated Parameters	α'	600.00	229.00
	ν'	4.4730	1.5350
Accident Reduction Factor	A.R.F.=((134.1503-149.0442)/134.1503)*100=-11.10 %		

increase was not significant. The increase was probably associated with two factors that are correlated with traffic accidents. First, the improvement of roadway conditions probably encouraged drivers to increase their operational speed, and second, wedge and level improvements applied to a limited area of the pavement may not have improved pavement skid resistance.

In reviewing the available literature, no reference was found that reported the safety impact of the wedge and level improvement as a single activity. However, wedge and level can be implicitly included under patching activities. If so, Creasey and Agent (*10*), on the basis of a literature review, indicated that resurfacing, patching, drainage, and deslick improvements reduced the expected level of traffic accidents by 16 percent.

Results of analyses performed in this study indicated that levels of traffic accidents (expected number of accidents and rate of accidents) increased after resurfacing improvements, but the increases were not significant at 5 percent probability level. One probable explanation for the increase is that resurfacing improved the quality of driving; therefore, traffic operational speeds increased, and the resurfacing impact on the level of accidents could not be compensated for by a small improvement in the skid resistance. The soundness of the result can be judged by comparisons with other available studies. Creasey and Agent (*10*) developed accident reduction factors for various highway improvements in Kentucky. Their recommendation was based primarily on engineering judgment and some before and after evidence. In another study (*11*), an empirical relationship was developed between the wet-pavement accident rate and skid resistance. The relationship indicated that the wet-weather accident rates drop for all rural highway classes when the pavement skid resistance increased. But on the basis of data from the resurfacing, restoration, and rehabilitation program, Brinkman (*12*) reported that resurfacing did not have a significant effect on the mean skid number of the tested sections selected in the study. In addition, on the basis of data from the same program, Tignor and Lindley (*13*) analyzed information from nine states and concluded that resurfacing increased the rate of accidents by about 2.2 percent, but the increase was not significant at 5 percent level. In the same study, Tignor and Lindley (*13*) found that resurfacing improvements resulted in an increase in the skid resistance of a 32-mi two-lane rural highway section in Alabama. The accident rate in this section increased by 11.85 percent after resurfacing, but the increase was not significant at 5 percent level. Thus, the results of the analyses performed in the present study are compatible with results obtained from others and a large data base for evaluating the effect of resurfacing improvements on the level of traffic accidents.

CONCLUSION

A methodology to evaluate the safety impact of highway improvements was developed. The evaluation was based on the expected number of accidents and expected accident rate, which were developed according to the Bayesian approach. The methodology was applied to evaluate the impact of wedge and level and resurfacing projects. The results indicated that wedge and level improvements did not have a significant effect on the level of the expected number of accidents or rate of accidents at 95 percent probability level on undivided rural roads having average daily traffic volume less than 4,000 vehicles per day. The results also indicated that resurfacing projects did not have any significant impact at 5 percent probability level on the level of traffic accidents on the same type of roads. The results were generally compatible with the information available in the literature.

REFERENCES

1. F. M. Council et al. *Accident Research Manual*. Report FHWA/RD80/106. FHWA, U.S. Department of Transportation, 1980.
2. C. C. Wright, C. R. Abess, and D. F. Jerrett. Estimating the Regression-to-Mean Effect Associated with Road Accident Black Spot Treatment: Towards a More Realistic Approach. *Accident Analysis and Prevention*, Vol. 20, No. 3, 1988, pp. 199–214.
3. C. V. Zegeer, D. W. Reinfurt, W. W. Hunter, J. Hummer, R. Stewart, and L. Herf. Accident Effects of Sideslope and Other Roadside Features on Two-Lane Roads. In *Transportation Research Record 1195*, TRB, National Research Council, Washington, D.C., 1988, pp. 33–34.
4. E. Hauer. On the Estimation of the Expected Number of Accidents. *Accident Analysis and Prevention*, Vol. 18, No. 1, 1986, pp. 1–12.
5. E. Hauer and B. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossing and the Safety Effects of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1987, pp. 131–140.
6. H. Al-Masaeid. *Bayesian Approach to the Estimation of Expected Number of Accidents*. Ph.D. thesis. Purdue University, West Lafayette, Ind., 1990.
7. I. Olkin, L. J. Gleser, and C. Derman. *Probability Models and Applications*, Macmillan Publishing Co., Inc., 1980.
8. C. Schmittlein and G. Morrison. Prediction of Future Random Events with the Condensed Negative Binomial Distribution. *Journal of the American Statistical Association*, Vol. 78, No. 382, June 1983, pp. 449–456.
9. J. Higle and J. Witkowski. Bayesian Identification of Hazardous Locations. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 24–36.
10. T. Creasey and K. R. Agent. *Development of Accident Reduction Factors*. Research Report UKTRP-85-6. Kentucky Transportation Research Program, March 1985.
11. R. Blackburn et al. *Effectiveness of Alternative Skid Reduction Measures*. Report FHWA-RD-79-21. FHWA, U.S. Department of Transportation, Nov. 1978.
12. C. Brinkman. Safety Studies Related to RRR Projects. *Transportation Engineering Journal of ASCE*, Vol. 108, No. TE4, July 1982, pp. 307–312.
13. S. C. Tignor and J. A. Lindley. Accident on Two-Lane Rural Highways Before and After Resurfacing. *Public Roads*, Vol. 44, No. 4, March 1981, pp. 137–139.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

On the Use of Accident or Conviction Counts To Trigger Action

E. HAUER, K. QUAYE, AND Z. LIU

The probabilistic properties of the process of identifying entities, such as drivers or intersections, for some form of remedial action when they experience N accidents within D units of time are explored. This mechanism for triggering action is referred to as an N-D trigger. On the basis of the probability distribution of the "time-to-trigger," it is concluded that in road safety the problem of false positives is severe, and therefore entities identified on the basis of accident or conviction counts should be subjected to further safety diagnosis. Moreover, the longer the N-D trigger is applied to a population, the less useful it becomes. The performance of the trigger depends on the choice of N and D, and guidance is offered on how best to choose them.

When at least five correctable accidents occur at an intersection within 12 months, the accident experience warrant for a traffic signal is satisfied (*1*, 4C-6). The same kind of warrant pertains to multiway stop signs (*1*, 2B-4). Similarly, it is common practice to flag a road section or intersection for selective enforcement, engineering study, or remedial action if the count of accidents occurring on it per unit of time or exposure exceeds a certain number. Also, a driver's license is suspended when the driver has accumulated and exceeded a certain number of demerit points, with points being erased from the driver's record a fixed time after conviction. All these situations have the following common conceptual structure:

> There exists a group of "entities," be they intersections, road sections, or drivers. On each entity occur events such as accidents or convictions. The process of event occurrence is characterized by a degree of randomness. If on an entity N or more events occur within a time (or distance or exposure) "window" that is of duration (or length or size) D, some action is triggered. The action may be the performance of an engineering study, increased enforcement, implementation of remedial measures, revocation of driving privileges, or the like. This method of triggering action will be referred to as the N-D trigger.

Because of the random nature of event occurrence, the N-D trigger has certain probabilistic properties that are of practical interest. For example, on occasion a flurry of accidents may occur at safer-than-average intersections. Whereas the accident warrant is thereby met and thus the installation of a higher-level traffic control device may be considered, in reality this is a false alarm. How often is the accident warrant met in this spurious manner? How does the frequency of false alarms depend on the choice of N and D? The same questions arise when remedial work is applied to falsely identified blackspots or when merely unlucky drivers lose their driving privileges.

The probabilistic properties of the N-D trigger have been explored earlier (*2*). Unfortunately, an error in one of the key equations makes all those results numerically incorrect. The purpose of this paper is to correct this error and to present results of practical interest to a broader audience.

Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada.

NOTATION AND ANALYSIS

Consider an entity on which "events" occur in accord with the Poisson probability law with a rate m per unit of time (or distance). When or where these events occur could be noted on a time or distance axis. Imagine a "window" of size D sliding along this axis. The time or distance from the origin until N events show in a window of size D for the first time is a random variable T. Thus T may denote the time when at an intersection a signal is warranted for the first time, it may designate the end of a road section that seems to be a blackspot, or T may stand for the instant when a driver has accumulated sufficient demerit points to warrant the suspension of the driver's license. The random variable T is the "time to trigger"; specific values of T will be denoted by t. The probability distribution $F_T(t)$ or the probability density function $F'_T(t) \equiv f_T(t)$ is needed to answer questions of practical interest.

It has been shown elsewhere (*2*) (a summary is given later) that in general,

$$F_T(t) = 1 - Ce^{-m\int_0^t p(\tau)d\tau} \tag{1}$$

where C is a constant of integration and $p(\tau)$ is the probability that there are $N - 1$ events in a window of size D when its right edge is on τ [the error mentioned earlier was to assume that $p(\tau)$ was independent of τ]. To find particular solutions and to determine what the constant of integration C is, initial conditions need to be specified. Two specific initial conditions are examined below.

Consider first an entity that has been in existence for some time. At time $\tau = 0$ when examination begins, fewer than N events are in the window extending from $-D$ to 0. For this, the "surviving entity case,"

$$F_T(t) = 1 - Ce^{-m\int_0^t p(\tau)d\tau} = 1 - e^{-m\bar{p}(t)t} \tag{2}$$

where $\bar{p}(t)$ is the mean value of $p(\tau)$ in $[0,t]$.

A good approximation to $F_T(t)$ is obtained when $\bar{p}(t)$ is replaced by the constant

$$\bar{p} = \frac{\dfrac{(mD)^{N-1}}{(N-1)!}}{\sum_{0}^{N-1} \dfrac{(\mathrm{mD})^i}{i!}} \times k \tag{3}$$

where

$$k = (0.0009 \times 0.5941^{-N} + 0.8754) + mD(0.0917 \times 1.3256^{-N} - 0.1273) + (mD)^2(0.0339 \times 1.4009^{-N} + 0.0073)$$

This approximation was developed by Liu (*3*) for values of N from 3 to 8 and various values of mD subject to the condition that $0 \leq \bar{p} \leq 1$. Typical values of $\bar{p}$ are shown in Figure 1.

Consider next an intersection or road that has been newly opened to traffic or a driver just licensed. We will call this the "new entity" case. Here,

$$F_T(t) = \begin{cases} 1 - \sum_{0}^{N-1} \dfrac{e^{-mt}(mt)^i}{i!} & \text{when } t \leq D \\ 1 - [1 - F_T(D)]\, e^{-\bar{p}(t)m(t-D)} & \text{when } t > D \end{cases} \tag{4}$$

where $\bar{p}(t)$ is the mean value of $p(\tau)$ in $[D,t]$. A good approximation is obtained when $\bar{p}(t)$ is replaced by $\bar{p}$. The nature of $F(t)$ is best illustrated by a numerical example.

Numerical Example 1: Consider an intersection, presently equipped with stop signs on the minor approaches, that has two correctable accidents per year as a long-term average ($m = 2$). Conversion to multiway stop control is warranted when five correctable accidents occur in a 12-month period (1, 2B-6). Thus $N = 5$ accidents and $D = 1$ year. The $F(t)$ curves corresponding to Equations 2 and 4 are shown in Figure 2.

Thus, if such an intersection was newly opened (or equipped), the probability that multiway stops will be warranted within 1 year is 5 percent (Point A); fully 73 percent of two-way stop-controlled intersections that in reality have the acceptable average of two correctable accidents per year will meet the accident warrant in the first 10 years of operation (Point B). If the intersection was in existence at time 0, the corresponding probabilities are 13 and 75 percent as shown by Points C and D.

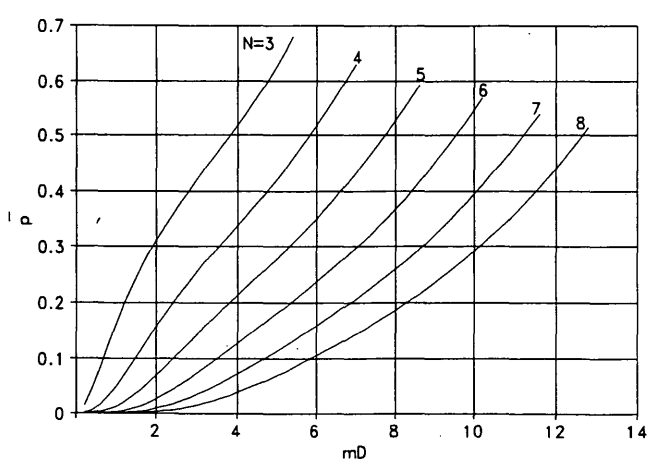


FIGURE 1 Values of $\bar{p}$ as a function of mD and N.

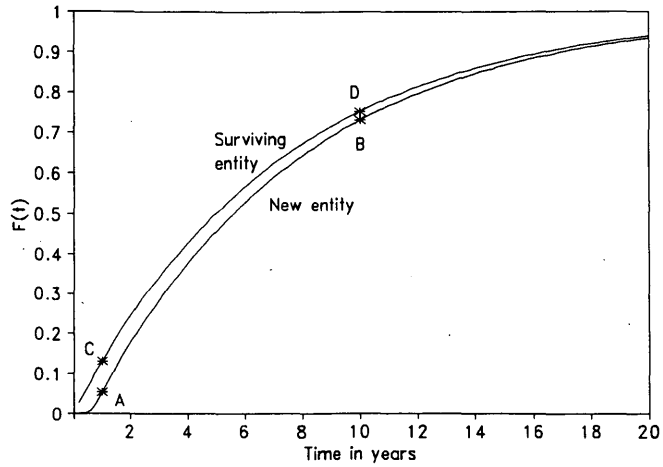


FIGURE 2 The probability distribution of the "time-to-trigger" when $mD = 2$ and $N = 5$.

A MICROCOSM

Having obtained the probability distribution of T, the functioning of the N-D trigger can be explored. It is planted into a simplified microcosm of professional practices that surround its use, and in this manner the workings of the trigger are illustrated.

Application to a Population

In Numerical Example 1 the m of the intersection was given. In reality, the N-D trigger is applied to a population of entities in which each entity has its own distinctive but unknown m. The purpose of the next example is to examine how an N-D trigger would perform if applied to such a population of entities.

Numerical Example 2: Consider a population of 1,000 intersections, all serving similar traffic flows. Assume that 900 have $m = 1$ accident per year, 90 have $m = 2$ accidents per year, and 10 have $m = 3$ accidents per year. It is not known which intersections have what m. Let $N = 5$ accidents and $D = 1$ year, as the MUTCD (*1*) recommends. At time = 0 no intersection had five or more accidents in the previous year. Table 1 indicates what is expected to happen by the end of the first year.

The entry in Column 5 is the product of the number of intersections in the population (Column 1) and the probability of an intersection to record five or more accidents before the end of the first year (Column 4). The hope was that the warrant will help to identify mainly the 10 "deviant" intersections that have an unusually high m. It turns out that in this particular population, the warrant may be expected to

TABLE 1 Performance of the 5–1 Trigger in the First Year

1 Number of Intersections	2 m (accidents/year	3 $\bar{p}$ eqn. 3	4 $F_T(1)$ eqn. 2	5 Number expected to meet warrant in first year
900	1.0	0.0122	0.0121	11.0
90	2.0	0.0697	0.1301	11.7
10	3.0	0.1433	0.3497	3.5
1000				26.2

identify during the first year nearly one-third of these. Thus, the number of "correct positives" in the first year is expected to be 3.5. This leaves the remaining 6.5 deviants unidentified in the first year—these are the "false negatives." Of the 26.2 intersections at which the warrant is expected to be met, 11 are in fact safer than average and do not deserve attention or treatment—these may be called the "false positives."

The main merit of this illustration is in showing how false positives and false negatives arise when an *N-D* trigger is applied to some real population. By definition, "deviant" is what constitutes a small minority, whereas "normal" is always the large majority. Even though only 1.2 percent of the 900 normal intersections will spuriously meet the warrant in the first year, because they are many, the number caught by the 5-1 trigger (MUTCD warrant) is still substantial. The upshot is that of the 26.2 intersections at which the warrant is expected to be met, we really wanted to catch only 3.5.

Whether a clever choice of N and D can improve the performance of the warrant will be examined later. At this point we merely note that the problem of false positives and false negatives is inherent in the *N-D* trigger. Furthermore, because normal entities must be many, the number of false positives is bound to be large.

Depletion

The discussion in Numerical Example 2 pertains to the use of the *N-D* trigger in a population of entities for 1 year. In practice, however, demerit point systems or blackspot identification procedures are being applied continuously. It is therefore important to examine the performance of the trigger over time.

Numerical Example 3: In Numerical Example 2, of the 900 better-than-average intersections with $m = 1$ accidents per year, $900 - 11.0 = 889.0$ are expected to survive 1 year. Of these, $889.0 \times 0.0122 = 10.8$ are expected to meet the warrant during the second year, and so on. Table 2 gives the number of intersections identified for inspection by the 5-1 trigger year after year.

The number of safe intersections (those with $m = 1$) expected to meet the warrant in consecutive years is 11.0, 10.8, 10.7, 10.6, 10.4, and so on. This group is a steady source of false positives for a long time. Of the 10 deviant intersections with $m = 3$, the number expected to meet the warrant in consecutive years is 3.5, 2.3, 1.5, 1.0, 0.6, and so on. Thus, in the course of a few years most of the truly deviant intersections would have been identified, and continued application of the *N-D* warrant becomes useless. Whereas in the first year 13 percent of the identified sites are expected to have $m = 3$, in Year 5 this declines to only 3 percent.

These are the main features of the process that arises when an *N-D* warrant is applied to a population of entities the m's of which change slowly. First, since normal entities are by definition many, this group supplies false positives at a nearly constant rate for a long time. Second, because deviant entities are by definition few and the *N-D* warrant is relatively good at catching them, their supply in the population is depleted in short order. Therefore, continued application of the warrant is bound to be unproductive.

Two practical conclusions can now be articulated. First, because an entity identified by the *N-D* warrant has a good chance of being a false positive, it is essential that it be subjected to a detailed diagnostic examination before any costly action is taken. To the extent that defensible procedures for such diagnostic examination are not yet part of the professional repertoire, they need to be developed.

Second, continued application of an *N-D* trigger to a population of entities that remains essentially unchanged is of dubious merit. It gives rise to a process by which some entities are always caught in the net, but ever fewer of these are of real interest. Thus while action (more traffic control, treatment of additional blackspots, revocation of licenses) continues apace, the entities that are its subject approximate an almost random selection from the general population. It seems that if the aim is to identify entities with unusually high m, the *N-D* trigger in the form of accident warrants or demerit points should be applied only to new or changed entities and always for a limited period of time.

Regression to the Mean

When remedial action is based on the *N-D* trigger, there is the danger of illusory success due to "regression to the mean."

TABLE 2 Performance of the 5–1 Trigger in the Course of 5 Years

m accidents/year	Expected number of entities identified				
	year 1	year 2	year 3	year 4	year 5
1	11.0	10.8	10.7	10.6	10.4
2	11.7	10.2	8.9	7.7	6.7
3	3.5	2.3	1.5	1.0	0.6
	26.2	23.3	21.1	19.3	17.8

This should by now be well known [even though reports ignoring it are still being published (*4,5*)]. The earlier numerical examples provide the setting for a particularly clear illustration of how this bias comes about.

Numerical Example 4: Imagine that traffic control has indeed been upgraded at the 26 intersections where the 5-1 warrant was met (see Table 1). At some later time, the local authority conducts a naive before-and-after study to assess the effect of this upgrading on safety. Assuming that the change in traffic control did not affect the m's and thus did not change the safety of these intersections, what are the conclusions of such a study likely to be?

Since $N = 5$, each converted intersection has had five correctable accidents in the year before conversion. If the m's of these sites did not change, reading from Table 1, we should expect to find after conversion $(1 \times 11.0 + 2 \times 11.7 + 3 \times 3.5)/26.2 = 1.71$ correctable accidents per intersection in a year. Thus, whereas there was no change in the m's, a simple before-and-after study is expected to indicate an impressive reduction from 5 to 1.71 accidents per intersection. This "improvement" is usually illusory because, by assumption, the long-term average number of accidents has not changed.

Thus the illusion of success due to "regression to the mean" in naive before-after studies is inherent in the automatic application of accident warrants that are of the *N-D* type. The hope is that accident warrants are not applied automatically. However, one application that is decidedly automatic occurs when demerit points are given to drivers who are convicted for violations of the highway traffic act. Typically, a set of actions (warning letter, interview, mandatory retraining, license suspension) is automatically triggered at some preset count of demerit points. This setting is explored next.

How Long Before You Lose Your License?

Most "good drivers" think it grossly unfair that one's driving privileges should be removed because of a run of bad luck. For a system of demerit points to be acceptable, the large majority of drivers whose licenses are revoked should have a genuinely high rate of convictions for the violation of traffic laws. Unfortunately, as will be shown, to devise such a system on the basis of convictions is nearly impossible. Just as in the case of intersection accidents, most drivers of a population who reach a predetermined count of convictions are false positives.

Rather than treading over old ground, by calculating the number of false positives for a typical population, the issue will be illumined from a different angle. The chosen tool of inquiry will be the "mean-time-to-trigger."

Consider the $F_T(t)$ for the surviving entity case (Equation 2). Integrating $tF'_T(t)dt$ from zero to infinity, it can be shown that the mean-time-to-trigger is

$$E\{T\} = \frac{1}{\bar{p}m} \tag{5}$$

This result is sufficient for a numerical example.

Numerical Example 5: In Ontario a driver's license is suspended when a person accumulates 15 demerit points. The number of demerit points depends on the type of offense and varies from 2 to 7. Demerit points are stricken from a person's record after 2 years ($D = 2$). With some 5 million drivers and 700,000 pointable convictions per year, the average Ontario driver receives about 0.14 pointable convictions per year. If the average number of points per conviction is 4, it takes about 4 pointable convictions in a 2-year window to reach suspension level.

Consider a better-than-average driver with $m = 0.1$ convictions per year. For this person, $\bar{p} = 0.000943$ (as calculated with $m = 0.1$, $N = 4$, and $D = 2$). For such drivers $E\{T\}$ is about 10,600 years. During a 50-year driving career, 1 of 212 such better-than-average drivers will have the license suspended.

Consider now another driver with a long-term conviction rate 5 times the population average ($m = 0.14 \times 5 = 0.7$ convictions per year). For this person, $E\{T\} = 15$ years. Thus, drivers of this kind will drive, on the average, 15 years before their license is suspended for the first time.

This kind of procedure may not be thought very satisfactory. First, there are millions of "better-than-average" drivers in Ontario. One in 212 such drivers will have the license suspended at some time. This seems unfair. Second, drivers who truly have an unusually large conviction rate may drive for a very long time before detection. The question arises whether it is possible to choose values of N and D to improve the performance of the trigger.

This question is best examined in light of the graphs in Figure 3, which is based on Equation 4. Point A pertains to the better-than-average drivers whose $m = 0.1$ convictions per year. With $N = 4$ convictions and $D = 2$ years, their $E\{T\} \simeq 10{,}600$ years. Point B describes the drivers whose long-term conviction rate is 5 times the population average ($m = 0.7$ convictions per year). With the same N and D, for these drivers $E\{T\} \simeq 15$ years.

If N is increased from 4 to 5, A moves to A′ and B moves to B′. This indeed decreases the probability of suspending the license of a better-than-average driver by a factor of 20. Unfortunately, this move further undermines our ability to detect bad drivers, tripling their $E\{T\}$ to 46 years.

The other option is to change D. If D is made 1 year instead of 2 while N remains 4, Point A moves to A″ and B to B″. For the better-than-average driver $E\{T\}$ is now 152,000 years, a shift in the desired direction. However the $E\{T\}$ for the

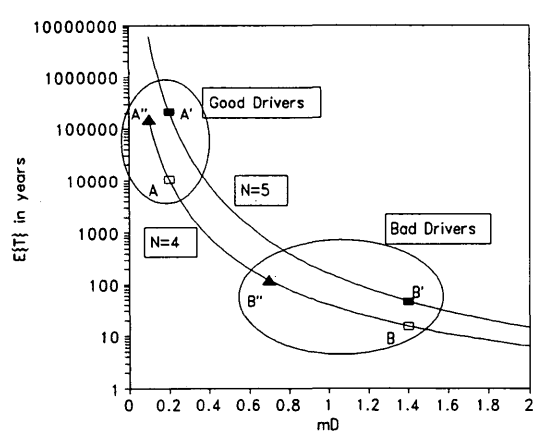


FIGURE 3 How $E\{T\}$ depends on N and D.

deviant drivers increases from 15 to 122 years. This is not what we wished to accomplish.

Thus, if one wishes to suspend the license of fewer better-than-average drivers, it will take even longer to catch those who have a truly deviant conviction rate and vice versa. Even though the performance of the *N-D* trigger in the demerit point context is rather dismal, there may be situations in which its performance is acceptable. Therefore it is still of interest to examine which choices of *N-D* are better than others.

CHOOSING THE BEST *N-D* PAIR

The main aim of the trigger is to identify entities having an unusually high m (m_{high}) and to do so, on the average, within a short time $E\{T|m_{\text{high}}\}$. There are many *N-D* pairs that all give the same $E\{T|m_{\text{high}}\}$. We propose to choose the *N-D* pair that makes $E\{T|m_{\text{low}}\}$ as long as possible for those entities having a relatively low m. This guidance for choosing an optimal *N-D* pair is applicable to the "new entity case" only. Finding an optimal *N-D* for the surviving entity case on the basis of the preceding proposition is meaningless since its structure suggests that one could always increase $E\{T|m_{\text{low}}\}$ by increasing D.

For a given m, N, and D, it can be shown that for the new entity case,

$$E\{T\} = [1 - F_T(D)]\left(D + \frac{1}{\bar{p}m}\right) + \frac{N}{m}\left(1 - \sum_{i=0}^{N} \frac{(mD)^i e^{-mD}}{i!}\right) \tag{6}$$

The way in which N and D influence $E\{T\}$ is shown in Figure 4. The figure shows $E\{T|m = 5\}$ for values of N equal to 10 and 15 and various values of D.

A number of observations can be made about $E\{T|m\}$ from this figure:

- $E\{T|m\}$ is generally high for small values of D;
- For any value of N, $E\{T|m\}$ has a minimum value, $E_{\min}\{T|m\}$;
- For large values of D, $E\{T|m\}$ approaches an asymptotic value of N/m; and
- $E_{\min}\{T|m\}$ increases as N increases.

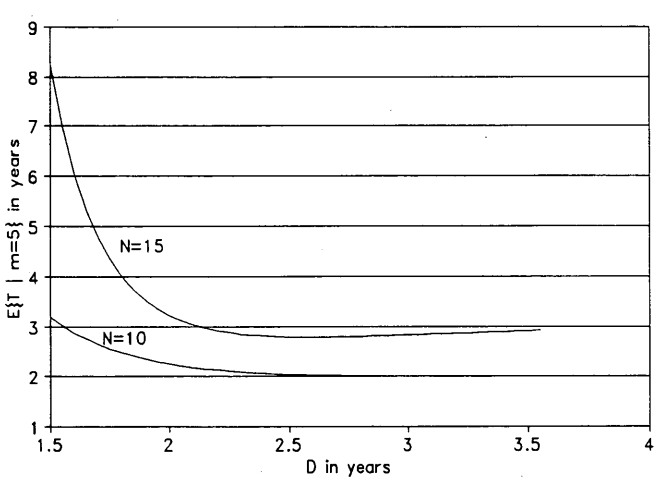


FIGURE 4 $E\{T|m = 5\}$ **as a function of** N **and** D **for the new entity case.**

As stated earlier, different *N-D* pairs can be found that yield the same values of a prespecified $E\{T|m_{\text{high}}\}$. However, from the foregoing, it is observed that the feasible values of N are those that meet the condition that the specified $E\{T|m_{\text{high}}\}$ is greater than or equal to $E_{\min}\{T|m_{\text{high}}\}$. This defines an upper bound for N. How an optimal *N-D* can be found will be illustrated with the following numerical example.

Numerical Example 6: In Numerical Example 2, of 1,000 similar intersections, 10 had a long-term average $m = 3$ accidents per year and 900 had $m = 1$ accident per year. Let these be the m_{high} and m_{low}, respectively. The aim is to identify the dangerous intersections within, say, 2 years. Several combinations of N and D that all have $E\{T|m = 3\} = 2$ years are given in Table 3.

Thus, the best *N-D* pair in this case is $N = 4$ and $D = 0.83$ years, which is quite close to the MUTCD warrant ($N = 5$, $D = 1$). Whereas intersections with $m = 3$ will be detected, on the average, in 2 years, intersections with $m = 1$ will survive, on the average, 30 years. Stated differently, one of 30 intersections with $m = 1$ accident per year will meet the warrant per year.

If the aim was to identify intersections with, say, $m_{\text{high}} = 5$ accidents per year, the optimal trigger would be $N = 9$, $D = 1.9$, which is quite different from that recommended in the MUTCD. In this case $E\{T|m = 1\} = 2{,}000$ years. A computer program for determining the best *N-D* pair is available on request. Some results are given in Tables 4 and 5.

DETECTION OF JUMPS IN *m*

So far we have imagined mainly the circumstance in which the m of each entity remains nearly constant over time and the task is to identify entities having an unusually large m. In this case, the longer the history of event occurrence for each such entity, the better one can detect the sought-after entities. One must therefore ask what sense it makes to use only the information contained in a window of size D and to disregard the portion of the event history to the left of the window.

There are, however, circumstances in which the use of a relatively short D might have appeal. One such circumstance is when a large and perhaps sudden increase in m might occur at some unknown time and the aim is to recognize it as soon as possible. Another is when D denotes distance, not time, and the window is sliding along some stretch of road. Here

TABLE 3 Combinations of *N* and *D* Yielding $E\{T|m = 3\} = 2$ Years

N	D	$E\{T\|m=1\}$
2	0.08	15.22
3	0.37	24.81
4	***0.83***	***29.85***
5	1.53	27.38
6	3.60	11.43

TABLE 4 Values of N, D, and $E\{T|m_{low}\}$ When $E\{T|m_{high}\}$ = 2 Years

m_{high}	m_{low}	N	D	$E\{T\|m_{low}\}$
3	1.5	4	0.83	9.31e+00
3	0.75	5	1.52	7.60e+01
3	0.3	5	1.52	3.47e+03
4	2.0	5	0.80	1.23e+01
4	1.0	6	1.23	1.88e+02
4	0.4	7	1.79	3.28e+04
5	2.5	6	0.80	1.58e+01
5	1.25	8	1.47	4.79e+02
5	0.5	9	1.89	3.76e+05
6	3.0	8	1.29	2.05e+01
6	1.5	11	1.82	1.84e+03
6	0.6	11	1.82	7.28e+06

the aim is to recognize when on a short stretch of road m is inordinately high.

It is unlikely that a trigger that makes no use of data outside the window of length or duration D can be an efficient device even in these circumstances. However, to give constructive advice on this matter is not simple. What can be shown with the results obtained so far is that in road safety the detection of jumps in m is a difficult task. In examining this question consider again the surviving entity case for which $E\{T\} = 1/(\bar{p}m)$ (see Equation 5).

In Numerical Example 5 a better-than-average Ontario driver was said to have 0.10 pointable convictions per year. To suspend the license of such a driver relatively rarely, $N = 4$ and $D = 2$ were chosen to make $E\{T|m = 0.1\} = 10{,}600$ years. Imagine that for some reason the m of this driver increases sevenfold to 0.7 convictions per year and the aim is to detect this jump not long after it occurs. For our purpose the instant of the jump in m can be taken as the time origin of a "surviving entity case." As was shown in Numerical Example 5, it will be 15 years on the average before the jump is detected. So, even with a relatively short D, a big jump in m still takes a long time to detect.

Similarly, imagine that the aim is to detect intersections where the m has jumped from 2 to 4. Using the best trigger ($N = 5$, $D = 0.8$ years, Table 4), it will still be 2 years on the average before the jump is recognized. Furthermore, in 1 out of every 12 intersections where no jump in m occurred (i.e., m remained two accidents per year), the trigger will be met.

In summary, when the m's are assumed not to change in time, the entire available historical record is relevant. To use the N-D trigger in this case makes little sense. The use of the N-D trigger might be justified when it is suspected that m can change suddenly and the aim is to detect such a change soon. Unfortunately, in the circumstances characteristic of road safety, the aim seems to be unattainable. The rate of event occurrence is too small to detect a sudden increase in m within few years of its occurrence without incurring the penalty of many false positives.

TABLE 5 Values of N, D, and $E\{T|m_{low}\}$ When $E\{T|m_{high}\}$ = 3 Years

m_{high}	m_{low}	N	D	$E\{T\|m_{low}\}$
3	1.5	5	0.96	2.09e+01
3	0.75	7	2.06	4.38e+02
3	0.3	8	2.80	1.57e+05
4	2	9	1.56	3.07e+01
4	1	11	2.73	2.76e+03
4	0.4	11	2.73	1.09e+07
5	2.5	29	2.71	1.78e+04
5	1.25	29	2.71	3.23e+12
5	0.5	29	2.71	1.47e+23
6	3	41	2.738	1.56e+08
6	1.5	41	2.738	5.65e+18
6	0.6	41	2.738	9.94e+33

SUMMARY

Procedures of the N-D trigger kind seem to be used in accident warrants and demerit point systems. The aim is usually to identify entities that are unsafe. In this paper the statistical properties of the N-D trigger are established and their repercussions for practical use are illustrated.

When the trigger is applied to a population of entities, inevitably there are false positives and false negatives. Since, by definition, normal entities are many and deviants are few, the trigger is often pulled unnecessarily. This leads to the conclusion that entities identified by the N-D trigger should be subsequently subjected to a sound diagnostic procedure capable of separating the wheat from the chaff.

The normal use of the N-D trigger is to scan the population of entities continuously or periodically. It is shown that if the unsafety of entities in the population changes slowly, repeated use of the trigger will deplete the supply of deviants while the majority of normal entities will provide a steady supply of false positives. After a few applications, the usefulness of the trigger may be exhausted.

A simple way to describe the performance of a trigger is by stating its mean-time-to-trigger as a function of the m of entities. Thus, for example, if a typical demerit point system is applied to a population of drivers, it takes 15 years on the average to remove the license of a driver whose conviction rate is 5 times the normal, and still 1 in 212 better-than-average drivers will lose the driver's license during the driving career.

If an N-D trigger is to be used, one may want to know how to choose the N-D pair sensibly. The choice can be made by first specifying the desired mean-time-to-trigger for entities considered deviant and then making the mean-time-to-trigger for a better-than-average entity as long as possible.

Whether to use an N-D trigger at all is an open question. It seems contrary to plain sense to disregard events that occurred outside the window of size D even if the intent is to detect sudden and large jumps in the rate of event occurrence. In any case, one can show that for the events of interest

(accidents at intersections or road sections, convictions for drivers) the rate of event occurrence is too small to have a realistic hope that sudden jumps can be detected soon after occurrence. It appears that the *N-D* trigger is not a particularly good device for identifying entities meriting attention.

DERIVATION OF $F_T(t)$

The equations in this section are based on the notation and description of the *N-D* trigger as given earlier. The probability $F'_T(t)\Delta t$ that N events show in a window of size D, for the first time, as its right edge moves from t to $t + \Delta t$ is the probability of the conjunction of the following three occurrences:

- The probability that N events did not materialize in a window of size D from when its right edge was at the origin until its right edge reached t [this is $1 - F'_T(t)$];
- The probability, $p(t)$, that when the right edge was on t, there were exactly $N - 1$ events in the window, given that the number of events in the window can be 0, 1, 2, . . ., $N - 1$; and
- The probability that an event was added to the window as the right edge moved from t to $t + \Delta t$ while none was deleted at the left edge (this is approximately $m\Delta t$).

The product of these three constituent probabilities is

$$F'_T(t)\,\Delta t = [1 - F'_T(t)] \times p(t) \times m \times \Delta t \tag{7}$$

When Δt approaches 0 this relationship turns into the simple separable differential equation $F'_T(t)/[1 - F'_T(t)] = p(t)m$, for which the general solution is

$$F_T(t) = 1 - Ce^{-m\int_0^t p(\tau)d\tau} = 1 - Ce^{-m\bar{p}t} \tag{8}$$

where C is a constant of integration and $\bar{p}(t)$ is the mean value of $p(\tau)$ in $[0,t]$. Elsewhere (*2*) it was assumed that $\bar{p}(t)$ was independent of t and was equal to the conditional probability

$$\bar{p}(t) = \frac{\dfrac{(mD)^{N-1}}{(N-1)!}}{\displaystyle\sum_0^{N-1} \frac{(mD)^i}{i!}} \tag{9}$$

This was an erroneous assumption since $\bar{p}(t)$ is in fact dependent on t.

Liu (*3*) developed correction factors for the expression for $\bar{p}(t)$ given in Equation 9. This was done by a simulation of the *N-D* trigger using various values of m, N, and D. Through an elaborate function fitting process, Liu found that a good approximation for $\bar{p}(t)$ is obtained when the expression in Equation 9 is multiplied by a correction factor k, where

$$\begin{aligned} k = {} & (0.0009 \times 0.5941^{-N} + 0.8754) \\ & + mD(0.0917 \times 1.3256^{-N} - 0.1273) \\ & + (mD)^2(0.0339 \times 1.4009^{-N} + 0.0073) \end{aligned} \tag{10}$$

REFERENCES

1. *Manual on Uniform Traffic Control Devices for Streets and Highways.* Federal Highway Administration, U.S. Department of Transportation, 1988.
2. E. Hauer and K. Quaye. On the Use of Accident and Conviction Counts To Trigger Action. In *Transportation and Traffic Theory* (M. Koshi, ed.), Elsevier, 1990, pp. 153–172.
3. Z. Liu. *N-D Trigger Simulation.* Master's thesis. Department of Civil Engineering, University of Toronto, Toronto, Ontario, Canada, 1992.
4. R. P. Bhesania. Impact of Mast-Mounted Signal Heads on Accident Reduction. *ITE Journal,* Vol. 61, No. 10, 1991, pp. 25–30.
5. N. Lalani. Comprehensive Safety Program Produces Dramatic Results. *ITE Journal,* Vol. 61, No. 10, 1991, pp. 31–34.

DISCUSSION

PATRICK BUTLER
National Organization for Women, 1000 16th Street, N.W., Suite 700, Washington, D.C. 20036–5705

By means of hypothetical accident and conviction rates, the paper evaluates the use of random events to single out individual intersections or drivers from apparently homogeneous populations for remedial treatment. For the intersection calculations, the necessity for equal exposure to risk is explicitly recognized by specifying that the intersections are "all serving similar traffic flows." The question of the known wide range in kilometers of exposure among drivers is ignored, however, in the model for the distribution of traffic convictions among drivers. By focusing on identifying "good" and "bad" drivers through annual conviction rates, the paper supports the erroneous insurance notion that driving risk can be accurately assigned to individual cars on an annual basis. In fact, however, the risk of traffic convictions and accidents is zero when a car is parked and increases kilometer by kilometer as a car is driven. The significant, predictable effect on insurance costs for drivers, especially for low-kilometer drivers, of ignoring individual exposure will be examined through consideration of the paper's Numerical Example 5.

With reference to a conviction rate of 0.14 per year averaged over all Ontario drivers, Example 5 specifies for the sake of its calculations that drivers with a conviction rate of 0.1 per year are "good" and that drivers with a conviction rate of 0.7 per year are "bad" and should be subject to having their licenses suspended. The Poisson calculations of Example 5 show that the current method specifying four convictions in 2 years (and other number-year criteria as well) for license suspension is very inefficient because too few 0.7-rate drivers and too many 0.1-rate drivers would meet this criterion. Nonetheless, the paper encourages continued efforts "to detect bad drivers," who, it appears to suggest, are those with annual conviction rates of 0.7 (or more). Table 6 tests the capability of such annual rate criteria for measuring driving risk.

Table 6 gives the annual kilometer exposures and conviction rates per kilometer chosen for four hypothetical drivers to produce the annual conviction rates used in Numerical Example 5. The kilometer-traveled value assigned each driver is within the range of what many drivers and cars typically travel in a year.

TABLE 6 One Year's Experience for Four Hypothetical Drivers

Driver	Km traveled during year (1)	Class average convictions per 10^6 km (2)	Convictions per year (1) x (2)	Paper's driver assessment based on convictions-per-year rate
1	10,000	10	0.1	"good"
2	70,000	10	0.7	"bad"
3	100,000	7	0.7	"bad"
4	10,000	70	0.7	"bad"

The table shows that a wide range of driving-exposure and conviction-rate combinations can produce a given low or high annual conviction rate. Although the class conviction rates per kilometer assumed for Drivers 1, 2, and 3 are close to the Ontario average (8.4 per 1 million km in 1990), the 0.7 annual rate calculated for Drivers 2 and 3 makes them "bad" drivers and, according to the paper, deserving of license suspension if they could be identified. The idea that individual "good" and "bad" drivers can be defined on the basis of annual rates of convictions or accidents is demonstrably not valid.

It may seem that those qualifying as "bad" drivers because of high annual kilometers driven are disadvantaged by being more likely to receive convictions and license suspensions than the low-kilometer "good" drivers. This is not true in the insurance pricing context, however. Since premiums are charged at annual class rates that are little affected by kilometers driven, the more driving that is done, the less is paid per kilometer for on-the-road insurance protection. With the same coverage and in the same class, Driver 2 would pay one-seventh the per-kilometer rate for insurance that Driver 1 does. Viewed another way, Driver 1 would pay 7 years of premiums for the 70,000 km of exposure to risk that Driver 2 pays one premium for.

Although reliable records of kilometers driven are not routinely kept for individual drivers, the kilometers driven by individual cars are registered on their odometers. Legislation requiring conversion of automobile insurance class rates from dollars per car year to cents per car kilometer as an anti-price discrimination measure is under consideration in several states because low-kilometer drivers are forced under the present system to pay much more per kilometer for identical insurance protection than high-kilometer drivers with cars in the same class. As now, premium would continue to be paid in advance to keep insurance in force, with the odometer limit to the prepaid protection displayed on the car's insurance card. Illegal odometer tampering, detected at the insurance company's annual audit or during accident investigation, would automatically void the policy.

A major block to such reform, however, is insurers' specious proof that classification of cars based on claim and conviction records provides cost justification for providing discounts for cars with "good" drivers and surcharges for cars with "bad" drivers. The criteria for these classifications are all variants of the Ontario number-year criterion for license suspension, such as years since last claim for "good" driver discounts and "bad" driver surcharges tied to claims and convictions within the last 3 years. The surcharge classes consistently experience more claims per car than the discount classes, both in fact and as modeled through Poisson calculations (*1*).

Although accidents (and traffic convictions) are random events, cars driven more kilometers than the average for their price class are more exposed to chance of accident and will be overrepresented in the class minority that has accidents and convictions. It is not "bad" drivers but overrepresentation of higher-kilometer cars that produces the higher accident averages that insurers invoke to justify the surcharging. Unlucky lower-kilometer insureds, already overcharged at "good" driver class rates, are more heavily overcharged by "bad" driver surcharges.

REFERENCE

1. P. Butler and T. Butler. Driver Record: A Political Red Herring that Reveals the Basic Flaw in Automobile Insurance Pricing. *Journal of Insurance Regulation,* Vol. 8, No. 2, Dec. 1989, pp. 200–234.

AUTHORS' CLOSURE

The object of our Numerical Example 5 is to quantify the implication of using an *N-D* trigger to identify "deviant" drivers. We base our analysis on an *N-D* trigger similar to the one used in Ontario (i.e., one with $N = 4$ pointable convictions and $D = 2$ years). In the hypothetical population used in Numerical Example 5, we therefore used the conviction experience of drivers in Ontario as a basis for defining a good driver as someone with an expected value of 0.1 convictions per year. To further explore the performance of the *N-D* trigger, we define a bad driver for this population as one whose expected number of convictions per year is 7 times that of our good driver (i.e., 0.7 convictions per year).

The discussant seems to be uncomfortable with the notion that drivers can thus be classified as good or bad on the basis of their accident or conviction record. This is because driving exposure (henceforth referred to as exposure) as measured by the number of kilometers driven per year by various drivers may be vastly different.

For a population of drivers with vastly different exposure, defining safety as the expected number of accidents or convictions alone can be problematic and misleading. Nonetheless, a nonhomogeneous group of this nature could be divided into sets of homogeneous groups in which drivers have fairly similar exposure. Within each homogeneous subpopulation, it is expected that the long-run average of the number of accidents or convictions per year will vary from driver to driver. Thus conditional on exposure, one is still in a position to classify some drivers as good and others as bad on the basis of their accident or conviction records.

For instance, in the table provided by the discussant, Drivers 1 and 4 have identical exposures of 1000 km/year. If the number of convictions in the table is representative of the long-run conviction experience of these two types of drivers, it would be appropriate to say that Driver 4 is worse than Driver 1. Let us suppose that an analyst desires to use the driving record of a population of drivers similar to these two to identify the bad drivers in the group for some form of remedial action. If an approach such as the *N-D* trigger is

used, our study provides the analytical tool for examining the implication of using any chosen *N-D* pair.

Essentially, our goal is to provide the quantitative framework for examining the effect of identifying entities (drivers or intersections) with an *N-D* trigger. In the case of a population with different exposures, the entities should be subdivided into homogeneous groups before exploring different choices of *N* and *D*.

Since the setting of insurance premiums and the fairness of premiums are beyond the scope of our study, we reserve our comments on those issues. In closing, though, it is worthwhile to note that the discussant appears to favor the use of exposure-based accident or conviction rates (e.g., accidents per vehicle-kilometer) as a basis for defining good or bad drivers. However, such a measure has its drawbacks, especially in cases where the relationship between exposure and the expected number of accidents or convictions is nonlinear.

Publication of this paper sponsored by Task Force on Statistical Methods in Transportation.

Cost-Effective Driver Improvement Treatment in Pennsylvania

LOREN STAPLIN

A written examination developed as one level of a multitiered driver improvement pilot program administered by the Pennsylvania Department of Transportation (PennDOT) was found in a 1-year evaluation period to result in significant reductions in accident and violation involvement rates and a cost savings of at least $150,000 per year relative to prior interventions by PennDOT. The examination and accompanying handbook provided to operators at the 6-point level of negligence conveyed the message that unsafe driving behaviors resulting in accidents and convictions reflect inappropriate choices for particular traffic scenarios and communicated the importance of individuals choosing to avoid specific behaviors that increase accident risk through real-world examples of familiar problem driving situations. Treatment development, administration, and evaluation activities are described.

The development, implementation, and evaluation of a pilot driver improvement program activity in Pennsylvania are described. Treatments at three distinct levels of demonstrated negligence were included in the pilot program. Later, their effectiveness was evaluated in comparison with previously existing postlicensing control procedures administered by the Pennsylvania Department of Transportation (PennDOT).

At Level 1—the first conviction received by a driver with no prior points—the pilot treatment was a personalized version of a warning/advisory letter. The previous practice of PennDOT was to send out an impersonal notice by postcard to first-time offenders. At Level 2, defined as the first time a driver accumulates 6 points on his or her record, a special point examination (SPE) emphasized the understanding of safe driving practices and skills to avoid high-conflict situations, replacing the behind-the-wheel reexamination previously administered by the state police. For Level 3, denoted by the multiple incidence of 6 points on a driver's record, the existing practice of requiring offenders to attend a departmental hearing was augmented in the pilot program with an educational treatment for small groups based on a values clarification curriculum—"Decisions for Safe Driving." The procedures and findings reported in this paper pertain only to the Level 2 SPE.

There was a single, unifying principle in the development of the multitiered pilot program: that unsafe driving behaviors resulting in accidents and convictions reflect inappropriate choices for particular traffic situations. The often-inevitable consequences of such choices can therefore not be explained primarily as events that "just happen" because of external factors, but instead they must be seen as outcomes that are potentially avoidable if drivers' decision-making skills can be improved. For the SPE treatment at Level 2 in this program, a written examination and study guide were developed that reinforce the importance of individual drivers choosing to avoid specific behaviors that increase accident risk. This message was communicated through real-world examples of familiar problem driving situations, both in the study guide and in the format of the majority of test items on the SPE.

The Scientex Corporation, Human Factors Division, Brookview Corporate Center, Suite 3000, 1250 South Broad Street, Lansdale, Pa. 19446.

TREATMENT DEVELOPMENT

The existing point system in Pennsylvania provided the framework for development of the SPE treatment in this research. Whereas many different combinations of conviction types can result in the same negligent operator status, in the majority of cases 6 negligent driving points on an operator's record reflects convictions for two moving violations. The placement of an individual in treatment Level 2 (i.e., the SPE) followed from the initial accumulation of 6 points on an individual's record, with allowable treatment alternatives defined in Section 1538(a) of the Pennsylvania Vehicle Code under the authority to require a driver to attend a driver improvement school or undergo a special examination.

The objectives in treatment development were to promote Pennsylvania's driver improvement priorities using interventions—or related approaches—with which earlier studies have achieved positive results. The desired outcome of driver improvement efforts in Pennsylvania is to modify the behavior of negligent drivers to produce measurable gains on accepted traffic safety indicators (i.e., chargeable accidents and selected violations connoting hazardous driving practice).

As noted above, a common thread linking the various driver improvement pilot program interventions is the fundamental orientation toward behavior modification through attitude change rather than through more conventional educational approaches stressing knowledge of rules and regulations or the acquisition of vehicle handling and maneuver skills (e.g., generic defensive driving courses). Pennsylvania's approach emphasizes skill acquisition, but deficient decision skills as opposed to a lack of driving skills are viewed as most responsible for unsafe behaviors. This approach, which is based on cognitive-behavioral theory, emphasizes internal versus external attributions to explain events; in other words, negligent drivers are taught that their traffic violations are the result of their own choices—albeit unconscious ones (*1*). Development of content for the SPE was guided by this tactic of fostering attitude change, insofar as the offenders subject

to this intervention were presented with test and study materials reinforcing the premise of individual choice and responsibility.

The subtasks required to develop the SPE included (*a*) the preparation of a suitable notification letter to inform offenders of their requirement to complete the examination, (*b*) the development of the test content, (*c*) pretesting of the SPE for item comprehension by naive and low-literacy individuals and for examination reliability; (*d*) the development of a driver's handbook to be distributed to offenders when they received the exam requirement notification letter, (*e*) an English-Spanish translation of key examination materials, and (*f*) training of test proctors responsible for administering the examination. The English-Spanish translation and training of test proctors will not be addressed in this paper.

Notification Letter

Following entry of conviction data into the system in Harrisburg, drivers who reached the 6-point level of negligence during the pilot program and who were therefore required to take and pass the SPE were informed of this requirement via a notification letter from PennDOT. The letter conformed to guidelines established during development of the pilot program warning/advisory letter (*2*).

A letter rather than a card was developed for reasons of privacy, as well as the degree of personalization permitted by the letter's expanded format. To emphasize the element of personalization, the offender's name was included in the salutation, and a statement of the specific offenses, associated violation and conviction dates, and points assessed leading to the driver's present SPE status was incorporated into the body of the letter.

Information pertaining to the nature or specific content of the examination was restricted to the accompanying driver's handbook sent to offenders with the notification letter; the letter clearly identified the handbook as the source of this information. With regard to examination scheduling information, a date by which time the offender must have completed the examination was stated in the letter, but the offender was permitted to choose the exact time and place of testing that was most convenient. Finally, the penalty for noncompliance with the examination requirement—removal of the offender's driving privilege for an indefinite period (until the exam is successfully completed)—was highlighted.

The format of the notification letter is shown in Figure 1.

COMMONWEALTH OF PENNSYLVANIA
DEPARTMENT OF TRANSPORTATION
Bureau of Driver Licensing
Harrisburg, PA 17123

(Address) (Operator record number)

(Date of birth)

Dear Motorist:

Because of your recent violation of Section ___ of the Pennsylvania Vehicle Code on ___, 19 __, and conviction on ___, ___points have been assigned to your driving record. You are required to take and pass a special written examination within thirty (30) days of the mail date of this letter, or your license will be suspended. No extensions will be granted. After passing the examination, two (2) points will be removed from your record, if you pass within the 30 days. If you do not pass the examination on your first try, you may take it again, but not on the same day. A more complete explanation of the examination requirement plus a Study Guide is included in the Handbook which accompanies this letter.

This examination may be taken by appointment only. To schedule your appointment, call 1-800-XXX-XXXX between the hours of 8:00 A.M. and 4:00 P.M., and identify yourself by name and driver's license number to the PennDOT operator who answers. PennDOT will schedule your examination time and location to be as convenient to you as possible. To be admitted to take the examination, you must show this letter and your valid driver's license or other acceptable identification as described in the Handbook. If you show up at the examination location without these items, or at a different time than your appointment, you will not be admitted to take the examination.

It is a violation of the Vehicle Code for you to appear at your special point examination and to show another person's license or someone else to appear showing your license. In either case, both you and the other person will be prosecuted by the Pennsylvania State Police. If your license is currently suspended, this letter does not authorize you to drive to or from your examination location.

A Spanish language version of this letter and the handbook may be obtained by calling the same telephone number given above. (Para obtener ambos, esta carta y el manual, en Espanol Telefones (1-800-XXX-XXXX.)

Sincerely,

Director,
Bureau of Driver Licensing

ALWAYS wear your seat belt--it's the LAW!

FIGURE 1 Notification letter sent to Level 2 drivers required to take the SPE.

Examination Content and Support Materials

This section describes development of specific test items, preparation of alternative equivalent forms of the examination, and pretesting of the examination to measure its reliability.

Test items, set in the context of a series of real-world accident scenario descriptions, required yes/no (true/false) responses evaluating the evidence of "choice" behavior for each involved driver/vehicle in each included traffic conflict situation. The evaluations called for the recognition of unsafe driving practices that were attributed (according to study material and examples presented in the handbook mailed to offenders) to inappropriate driver decisions. The emphasis on driver decision making, in turn, underscored the message of individual choice and individual responsibility, which was fundamental to the ultimate objective of attitude change in PennDOT's approach to driver improvement.

To begin, the creation of test items for the SPE proceeded through a comprehensive review of material implemented in other jurisdictions and was targeted to a comparable (i.e., intermediate) level-of-negligence offender population. To identify all relevant sources and examples of candidate test items, a computer search through data bases including NTIS, TRIS, and PSYCINFO was performed, using the DIALOG access system. From this search, 116 reference citations were identified on the topic of driver improvement programs. The most relevant literature identified was borrowed from Northwestern University's Transportation Engineering Library.

The published record was supplemented by direct queries about research and program development efforts in progress in selected states. Telephone contacts were made with driver licensing officials in 16 states (Alabama, California, Delaware, Florida, Maine, Maryland, Michigan, Montana, North Carolina, North Dakota, Ohio, Oregon, Rhode Island, Washington, West Virginia, and Wisconsin) to identify and gather background information on driver improvement programs and, specifically, written examination materials. Sample copies of these exams were requested. From conversations with these agencies, it was found that very few of these states had driver improvement written examinations or examination material (or programs for that matter). The materials of greatest po-

tential use for developing the written examination in this task were obtained from Montana, North Carolina, Florida, and Washington.

Additional sources contacted to obtain candidate examination material for this subtask included the National Highway Traffic Safety Administration; safety-oriented organizations such as the National Public Services Research Institute, American Automobile Association, National Safety Council, and the Insurance Institute for Highway Safety; and university-affiliated centers for transportation research including the University of Michigan Transportation Research Institute and North Carolina's Transportation Research Center.

As candidate test items were compiled from these sources, an initial screening of content separated items pertaining principally to "rules and regulations" from those more generally addressing safe driving practices; items dealing with the effects of alcohol/controlled substances on driving performance and with vehicle restraint system facts and characteristics were included with the latter category. Items in the safe driving practices category that could be presented using diagrammatic information with a minimum amount of text were accorded the highest preference for inclusion in the SPE, together with information pertaining specifically to sanctions implemented in Pennsylvania.

On the basis of this information, a total of 20 items/responses were developed, with the objective of limiting the time needed to complete the examination to a maximum of 30 min.

Next, the candidate test items were sorted into categories according to content and judged level of difficulty per item. The judgments of difficulty were performed both by project staff and by drivers naive to this project. The outcome of this subtask was three alternative, equivalent test forms.

The content-sorting of items was differentiated among candidates according to specific areas of knowledge and specific driving situations. The knowledge of sanctions (i.e., the consequences of continued negligence) was stressed, in addition to an awareness and understanding of safe driving practices concerning visual search, speed control, and direction/maneuver control. Situational variables considered in the sorting of candidate test items included residential, urban arterial, and freeway driving conditions; night and other low-visibility conditions; roadway geometric variables including vertical and horizontal curvature, intersections, protected and unprotected turning situations, entry and exit ramps from limited-access highways, and high-speed merging-weaving situations; and railroad grade crossings.

The test items—other than those addressing knowledge of sanctions and related traffic safety issues—contained both text and diagram question elements. All questions were edited by literacy experts to ensure their suitability for administration to drivers at a sixth-grade reading level. A preliminary consultation with a literacy specialist addressed broad guidelines for presenting information of this nature in the most easily understood format. The bulk of the literacy screening effort followed preparation of a completed draft of the examination and support materials, however. At this time, the literacy specialist conducted a readability analysis to assess the level of the examination materials, then modified text as required to convey the desired information at the sixth-grade level. The modified text was returned for review by project staff to confirm that the meaning of questions was consistent with original intent.

The diagrams depicted real-world accident scenarios, identifying each contributing or involved driver/vehicle, pertinent roadway geometric features, traffic conditions, visibility and environmental factors, and any additional information helpful in an after-the-fact definition of appropriate and inappropriate driver decisions within a specific situation. Also, more than one test item was generated for each accident description and diagram. In general, test takers were required to discriminate those circumstances in which unsafe practices could be attributed to decisions and to further evaluate the consequences of those decisions as probable causes of the indicated conflicts/accidents diagrammed in the examination forms.

The three alternative, equivalent forms of the examination were prepared by altering the labeling of driver/vehicle question elements for a particular accident diagram and/or rephrasing the question so that the correct response changed from true to false or vice versa.

An example of an accident scenario and accompanying examination question is shown in Figure 2.

It was important to assess the reliability of the SPE before introducing its use to the driving public. If a measuring instrument is to be of any value, the results it produces should be highly consistent and reproducible—this is the operational definition of test reliability.

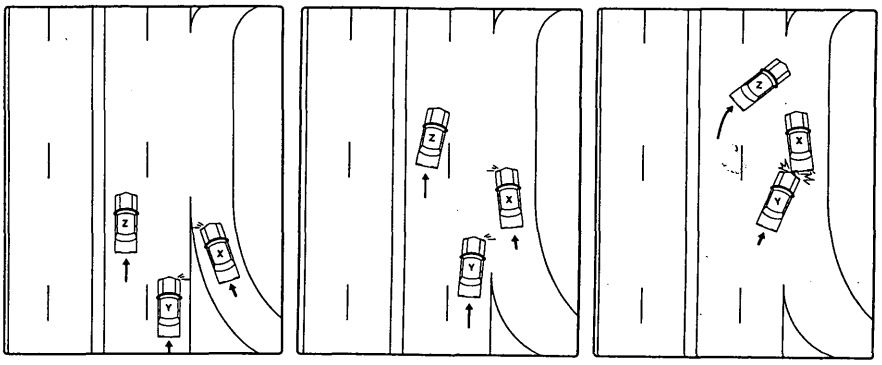
The subject sample selected for examination reliability pretesting consisted of 82 PennDOT personnel who were naive to this project. Each subject completed each of the three test versions; however, the order of completion of the versions was randomized, to rule out the effects of boredom/fatigue. An item analysis was conducted to consider significant asymmetries in the correct response rates for items in the "knowledge" and "accident situation" segments of the SPE and to isolate the probable sources of difference across versions of the examination. Items with correct response rates of less than 70 percent indicated either an unacceptable level of comprehension/retention of study material, exaggerated difficulty of test questions due to the number of included information elements, or awkward wording of the test question. Twenty-six items (across the three versions) therefore underwent revisions after review by the author and PennDOT project management personnel.

Driver's Handbook

As noted earlier, offenders notified by the department that they must complete the SPE received a handbook in conjunction with the notification letter. The handbook was designed to prepare drivers for the exam, specifically, as well as more generally to communicate the overall goals and structure of the driver improvement program and the dependence of safe driving outcomes on safe driving decisions. The handbook contained three sections: The Examination Requirement, Sample Test Items and Study Guide, and General Examination Information.

The initial section explained the examination requirement as it relates to the driver improvement program in Pennsylvania, including information describing the entire range of

Look at the accident diagrammed below, then answer the question at the bottom of the page as if you are Driver Z.



In the first diagram:	A couple of seconds later:	The accident occurs when:
Driver X is speeding up on an entrance ramp to the freeway, with his left turn signal flashing.	Driver X is about to enter the freeway.	Driver Z cuts in front of Driver X to exit the freeway.
Driver Y is traveling in the right lane of the freeway, with his right turn signal flashing.	Driver Y is just beginning to turn toward the exit ramp.	Driver X puts his brakes on hard to keep from hitting Driver Z.
Driver Z is traveling in the left lane of the freeway.	Driver Z begins a sharp turn from the left lane toward the exit ramp.	Driver Y crashes into the rear of Driver X.

QUESTION: Your decision to pass in front of Drivers X and Y to exit the freeway was responsible for the most unsafe act in this accident situation.

TRUE_____ FALSE_____

FIGURE 2 Example of SPE question addressing safe driving practices.

sanctions and remedial program activities triggered by different levels of demonstrated driving negligence. The link between driving negligence and faulty driver decision making was established in this section.

The Study Guide section described the examination's emphasis on knowledge of safe driving practices rather than memorization of traffic rules and regulations, then proceeded with a discussion of critical aspects of driving performance in specific driving situations. The discussion was designed to provide instruction regarding appropriate behavior for the same traffic conflict situations and (search, speed, and direction control) performance factors targeted in the development of the test items themselves. Following this discussion, example items were provided using the same traffic situation diagram approach designed for use in the examination, with each example including the correct answer and an explanation of how that answer is most consistent with the traffic safety lessons presented in the handbook. However, the range of all possible test item situations/questions was not covered in the handbook examples. Instead, apart from questions addressing knowledge of sanctions, examinees were required to generalize from the material presented in the handbook to the specific situation described in a given test item. This represents a clear departure from an approach in which offenders are merely required to memorize a set of facts to match with answer alternatives on an examination. The handbook very clearly communicates this difference to drivers who are required to complete this test.

The section labeled General Examination Information described the mechanics of completing the examination requirement. This included a listing of test sites, scheduling information, how and when drivers would receive official notice of their examination result and consequent license status, how to reschedule the examination if the first attempt resulted in failure, and a description of the examination protocol and expectations concerning behavior at the test site.

The complete handbook is presented as an appendix to the PennDOT final report completed for this project (*2*).

TREATMENT IMPLEMENTATION

The implementation of the SPE in the pilot program was carried out by trained examiners employed by PennDOT. The exam was administered over a 4-month period to groups of

no more than 18 examinees in each of 15 testing centers distributed across the commonwealth.

As noted earlier, the intent of this project was to compare the SPE with the procedure previously used by PennDOT as an intervention at the 6-point level of negligence, a behind-the-wheel (BTW) test of driving skill. The drivers receiving the SPE and the BTW test thus defined the treatment and comparison groups in this research, respectively. All eligible drivers received the SPE during the 4-month period assigned as the "treatment" interval in this project; the comparison interval was the preceding 4-month period, when all eligible drivers received the BTW examination procedure. The resulting composition of the treatment and comparison groups was subsequently screened to ensure that each was composed of unique sets of operators.

The total number of drivers included in the evaluation data set for the SPE was 19,194. The treatment group size was 11,291, or 58.83 percent of the total; 7,903 drivers, or 41.17 percent of the total, were included in the comparison group. Again, these groups were composed of all eligible drivers, statewide, whose conviction experience required completion of the PennDOT Level 2 intervention during the time period in question.

A summary of the age and gender composition of the treatment (T) and comparison (C) groups is as follows: average age_T = 28.7, average age_C = 28.6; age $range_T$ = 16–86, age $range_C$ = 16–83; percent $male_T$ = 78.8, percent $male_C$ = 79.0; and percent $female_T$ = 21.2, percent $female_C$ = 21.0. To help convey the similarity in group makeup, the complete distributions of driver ages for each group are shown in Figure 3.

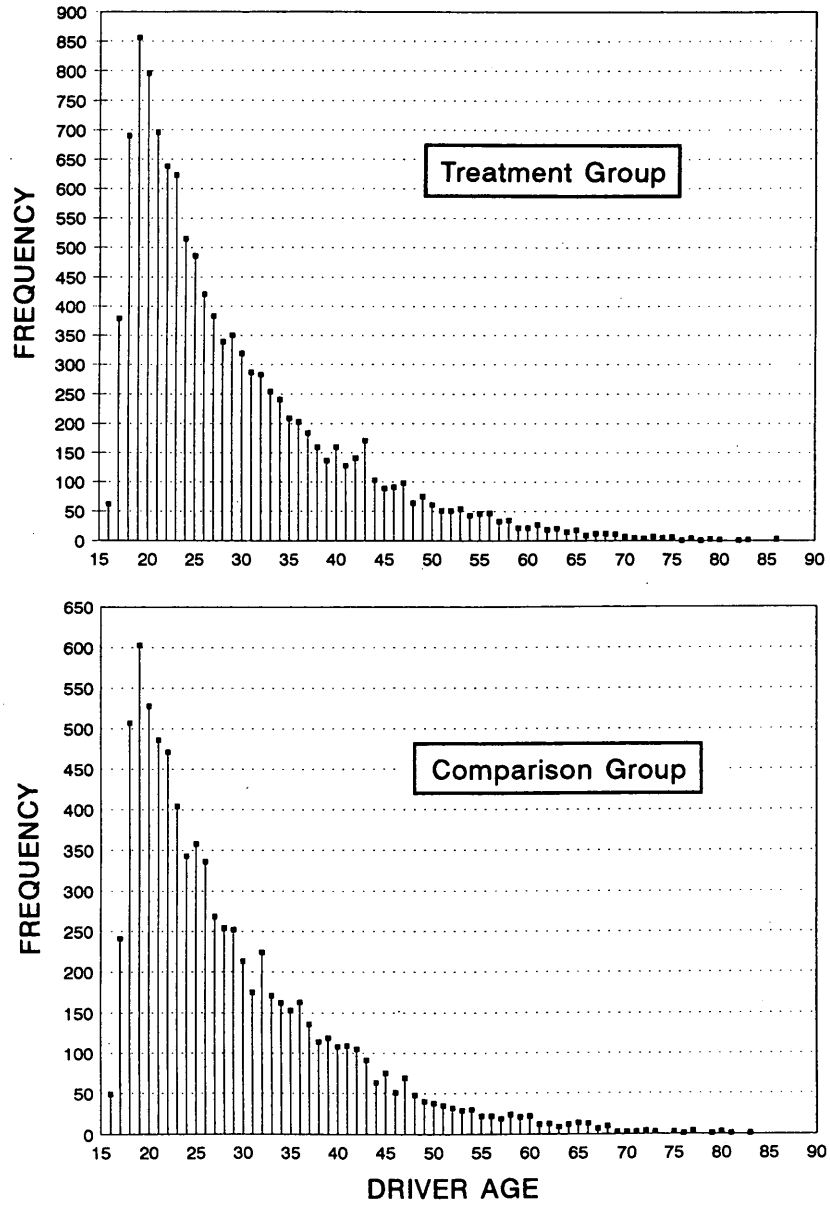


FIGURE 3 Distribution of driver ages for Level 2 treatment and comparison group drivers.

EVALUATION OF TREATMENT EFFECTIVENESS

The effectiveness of the SPE implemented in the pilot program was evaluated in terms of the subsequent accident and violation experience of the treatment and comparison group drivers relative to a specified reference date for each individual in each group.

The dependent measures used to gauge treatment effectiveness included the frequency of driver involvement in three accident and three violation categories. The accident categories included all chargeable accidents, single-vehicle accidents only, and multiple-vehicle (chargeable) accidents only. The violation categories included all point and major nonpoint violations combined, moving (point) violations only, and major nonpoint violations only. The violations included within each evaluation category are presented in Table 1.

Frequency counts for each of the dependent measures for the treatment and comparison groups were obtained at each program intervention level over variable amounts of time, depending on the date of implementation of a given treatment. The project schedule afforded 12-month evaluation periods for violation and for accident experience.

Two analytical procedures were applied to the evaluation data. First, the cumulative percentages of the treatment and comparison groups experiencing their first accident (or violation), on a month-by-month basis after the treatment (reference) date, were plotted. This type of plot is simply the inverse of the traditional "survival curve" that indicates the percentages of each group remaining accident (or violation) free over time. In this type of analysis, treatment effectiveness is demonstrated by a lower (cumulative) percentage of group members having experienced an accident (or violation), relative to the comparison group, at any specified time during the evaluation period.

Next, to determine the significance of observed differences in the relative accident (and violation) experience of the treatment and comparison groups, chi-square (X^2) tests were performed at planned milestones during the evaluation period for each intervention level. The comparisons were performed for the frequencies of accidents (and violations) observed at the 3-, 6-, 9-, and 12-month milestones. Only a single incident (accident or violation) was permitted for any given operator in the treatment and comparison groups for the X^2 tests; multiple incidents were excluded from this analysis of treatment effectiveness. (In fact, it was observed that the numbers of drivers with multiple event involvements at Level 2 in the pilot program were split almost exactly evenly between the treatment and comparison groups.)

TABLE 1 Allocation of Violation Codes to Violation Subgroups

Violation Code	Violation Subgroup	Violation Description
1543	Major NP	Driving while operating privilege is suspended or revoked
3112A3I	Moving Pt	Failure to stop for a red light
3114A1	Moving Pt	Failure to stop for a flashing red light
3302	Moving Pt	Failure to yield half of roadway to oncoming vehicle
3303	Moving Pt	Improper passing-overtaking driver to maintain speed; passing driver to pull in at safe distance
3304	Moving Pt	Improper passing on the right
3305	Moving Pt	Improper passing on the left - clear distance ahead
3306A1	Moving Pt	Improper passing on a hill
3306A2	Moving Pt	Improper passing at a railroad crossing or intersection
3307	Moving Pt	Improper passing in a no-passing zone
3310	Moving Pt	Following too closely
3321	Moving Pt	Failure to yield to driver on the right at intersection
3322	Moving Pt	Failure to yield to oncoming driver when making left turn
3323B	Moving Pt	Failure to stop for stop sign
3323C	Moving Pt	Failure to yield at yield sign
3324	Moving Pt	Failure to yield when entering or crossing roadway between intersections
3332	Moving Pt	Improper turning around - illegal U-turns
3341	Moving Pt	Failure to stop for flashing red lights or gate at railroad crossing
3344	Moving Pt	Failure to stop when entering from alley, driveway or building
3345A	Moving Pt	Failure to stop for school bus with flashing red lights
3361	Moving Pt	Driving too fast for conditions
3362	Moving Pt	Exceeding maximum speed
3365B	Moving Pt	Exceeding special speed limit in school zone
3365C	Moving Pt	Exceeding special speed limit for trucks on downgrades
3367	Major NP	Racing on highways
3542A	Moving Pt	Failing to yield right-of-way to pedestrians in crosswalks
3702	Moving Pt	Improper backing
3714	Moving Pt	Reckless driving
3731	Major NP	Driving under influence of alcohol or controlled substance
3733	Major NP	Fleeing or attempting to elude police officer
3734	Major NP	Driving without lights to avoid identification or arrest
3742	Major NP	Accidents involving death or personal injury
3743	Major NP	Accidents involving damage to attended vehicle or property

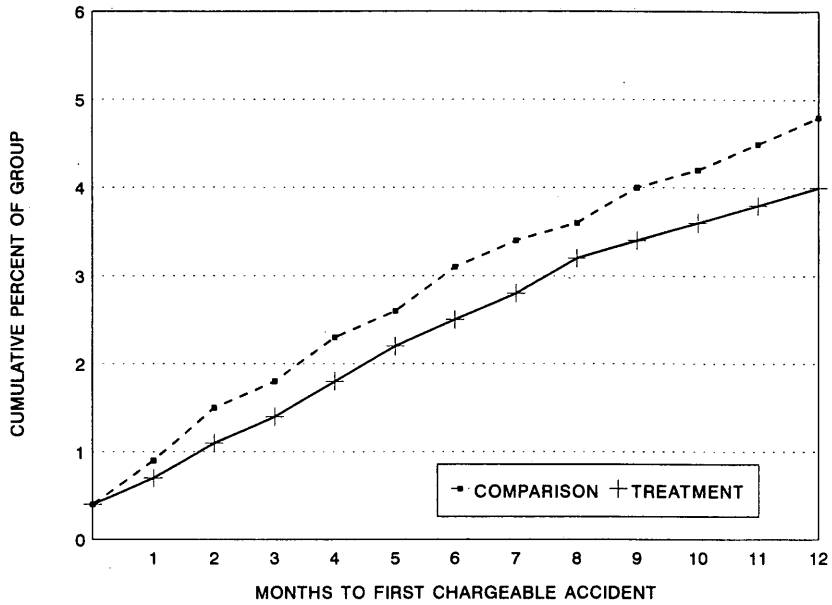


FIGURE 4 Cumulative percentages of Level 2 treatment and comparison group drivers with any chargeable accident involvement each month during the evaluation period.

Results of Analyses of Accident Experience

The curves plotted in Figure 4 describe a consistent benefit of the SPE versus the previously administered BTW exam procedure in terms of the overall chargeable accident experience of the drivers in the present study groups. The superior performance of the treatment group is evident across the entire evaluation period—12 full months; notably, this effect becomes more pronounced at greater intervals after the reference date for each driver.

The strongest difference between the treatment and comparison groups is apparent in the curves describing multiple-vehicle accident involvement (see Figure 5). A separation between the experience of the two evaluation groups that favors the treatment group drivers becomes evident at the end of the first month, and it widens consistently throughout

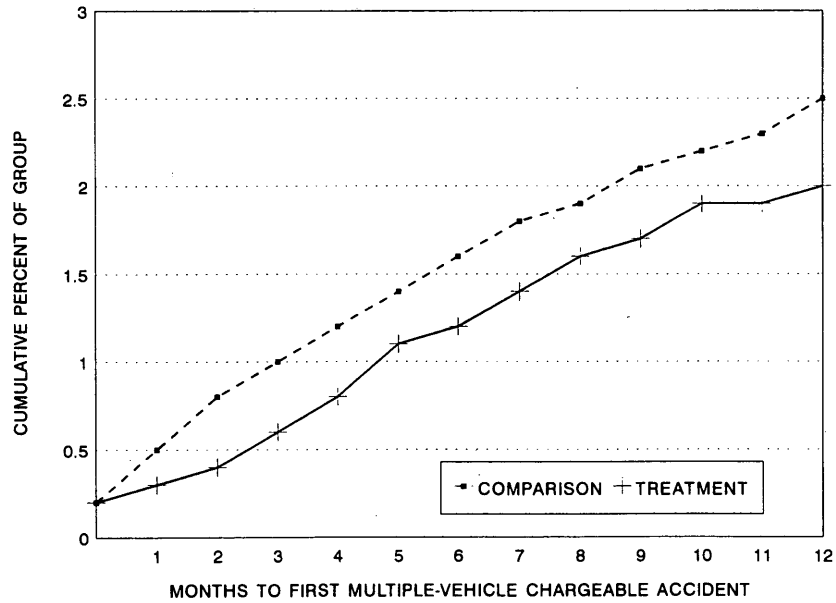


FIGURE 5 Cumulative percentages of Level 2 treatment and comparison group drivers with multiple-vehicle accident involvement each month during the evaluation period.

the year-long evaluation period. With respect to single-vehicle accidents, the experience of the treatment group drivers did not differ appreciably from the comparison group for the first few months, was only slightly (but consistently) superior from Months 3 through 9, then demonstrated a more pronounced benefit by Month 12 of the evaluation period, as shown in Figure 6.

X^2 tests performed on the accident data first indicated significant reductions in accident frequencies for treatment group drivers after 3 and 12 months for the all chargeable accident types category, whereas the reduction in accidents also approached significance at the 6-month ($p < .08$) and 9-month ($p < .09$) intervals. After a 3-month period, the treatment group drivers exhibited significantly fewer than expected accidents, whereas the comparison group drivers exhibited a significantly larger-than-expected number of accidents ($X^2 = 5.41$, df = 1, $p < .025$). After 12 months, the treatment group exhibited a much lower-than-expected frequency of all chargeable accident types, whereas the comparison group accident involvement was higher than expected. This result was significant at $p < .025$ ($X^2 = 5.64$, df = 1).

Next, the treatment group exhibited significantly better performance (fewer accidents) after 3 and 6 months for the multiple-vehicle accidents category, while approaching significant reductions in accidents at the 9-month ($p < .07$) and 12-month ($p < .08$) intervals as well. The superior performance for the treatment versus comparison group drivers was demonstrated by significant chi-square test results, described at the 3-month interval by $X^2 = 10.61$ (df = 1, $p < .005$) and at the 6-month interval by $X^2 = 4.83$ (df = 1, $p < .05$).

No significant differences were indicated when the dependent measure consisted of frequency counts of single-vehicle accidents, though Figure 6 shows a consistent pattern of accident reduction for the treatment group.

Results of Analyses of Violation Experience

As shown in Figure 7, the treatment group drivers experienced a reduced rate of convictions for violations of any type throughout the evaluation period relative to the comparison group drivers. The apparent benefit of the SPE examination was evident at a nearly constant level for a full 12 months.

The curves describing the experience of the treatment and comparison groups for moving/point violations during the evaluation period, as shown in Figure 8, almost exactly reproduce the pattern of results shown for all violations. The benefit of the treatment in reducing major nonpoint violation experience is less apparent, however, since both groups of drivers evidenced identical rates of (first) convictions of this sort at multiple points during the evaluation period (see Figure 9).

For the all violation types category the X^2 tests revealed a significant reduction in the number of violations experienced by the treatment group drivers versus the comparison group drivers at every interval up to and including 12 months from the reference date. These reductions in violation frequencies reached the $p < .005$ level of significance after 3 months ($X^2 = 10.85$, df = 1), 6 months ($X^2 = 12.80$, df = 1), and 12 months ($X^2 = 8.81$, df = 1); at the 9-month milestone the difference was significant only at $p < .05$ ($X^2 = 4.78$, df = 1). For each time period, the treatment group exhibited fewer than expected violation frequencies for the proportion of drivers in the treatment group and the total number of violations received by the sample.

Identical results were obtained for analyses considering the effectiveness of the SPE on frequency counts of moving/point violations only. The treatment group showed significantly fewer than expected moving/point violations at 3 months ($X^2 = 11.80$, df = 1, $p < .005$), 6 months ($X^2 = 11.24$, df = 1, p

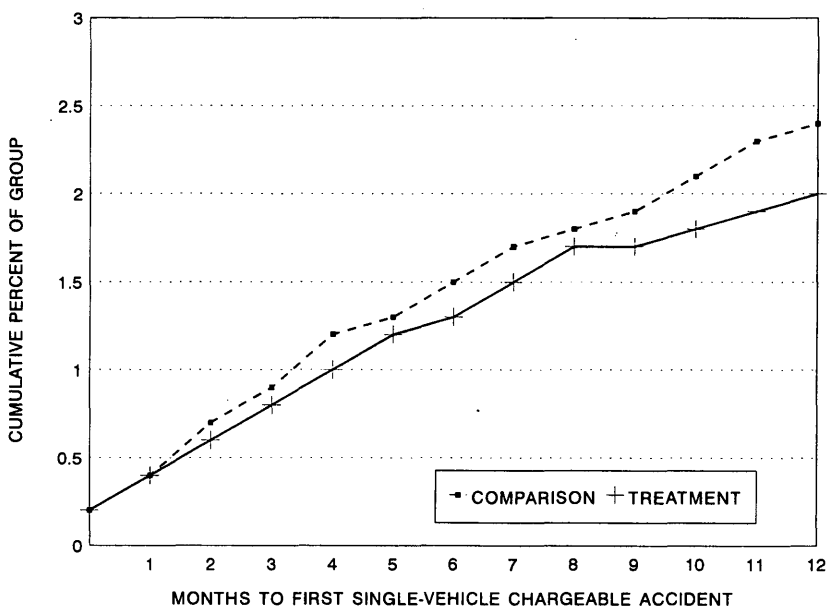


FIGURE 6 Cumulative percentages of Level 2 treatment and comparison group drivers with single-vehicle accident involvement each month during the evaluation period.

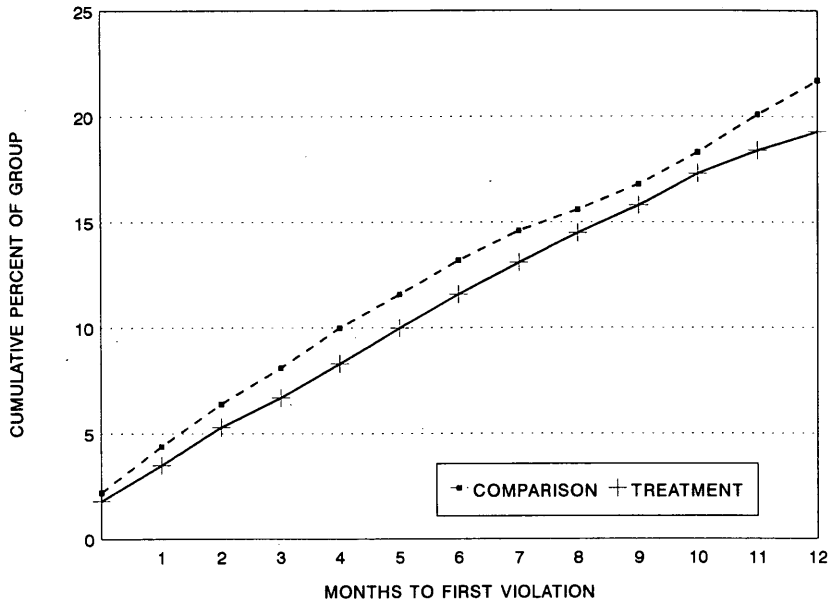


FIGURE 7 Cumulative percentages of Level 2 treatment and comparison group drivers with convictions for any violation type each month during the evaluation period.

< .005), 9 months (X^2 = 5.15, df = 1, p < .025), and 12 months (X^2 = 8.75, df = 1, p < .005).

Chi-square tests indicated that the SPE treatment had no significant effect on major nonpoint violation experience of the treatment group versus that of the comparison group, although modest gains in performance for the treatment group are apparent in Figure 9 at three of four evaluation milestones.

PROGRAM COST COMPARISON

Before implementation of the Level 2 SPE developed in this research, the equivalent of at least eight Pennsylvania State Police driver license examiners was required to administer the BTW reexamination. Through group administration of the new SPE treatment, personnel requirements for PennDOT

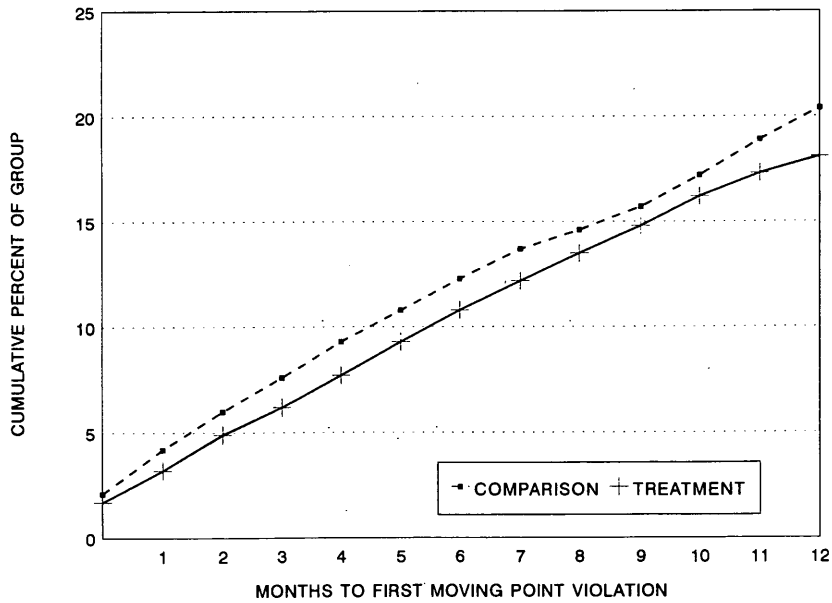


FIGURE 8 Cumulative percentages of Level 2 treatment and comparison group drivers for moving/point violations each month during the evaluation period.

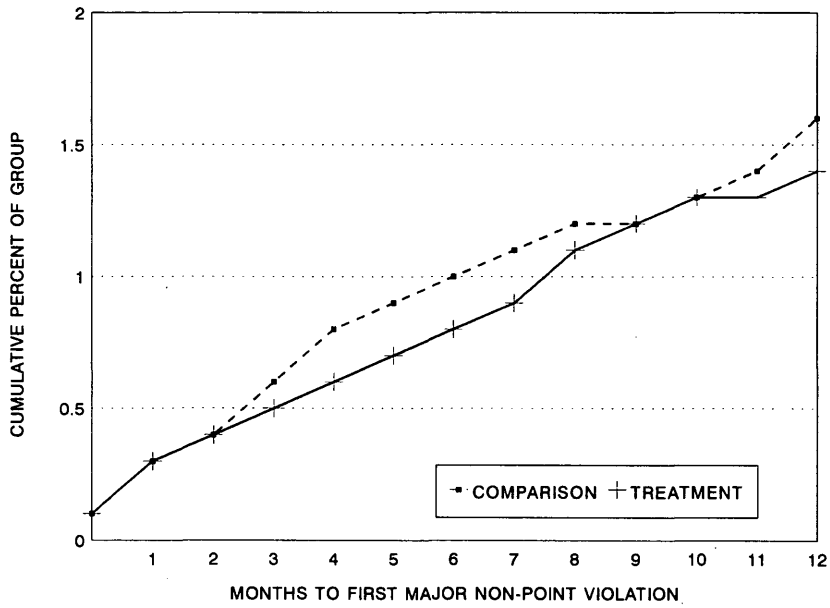


FIGURE 9 Cumulative percentages of Level 2 treatment and comparison group drivers for major nonpoint violations each month during the evaluation period.

were reduced to the equivalent of three driver safety examiners. The net savings to the commonwealth of this change has been conservatively estimated at greater than $150,000 per year by Department of Transportation officials. In the future, automated scoring and examination administration are projected to result in additional savings.

SUMMARY AND CONCLUSIONS

Using the results of related studies of postlicensing control activities in other states as a starting point, pilot program interventions were developed and implemented in this project including an SPE administered to drivers upon first reaching the 6-point level of negligence, most commonly resulting from two moving violations. This intervention stressed safe driving practices rather than knowledge of rules and regulations, and both the study materials distributed to drivers and the examination instrument itself communicated to each individual the responsibility for making safe driving decisions. Exhaustive administration of the novel exam treatment within a bounded (4-month) interval was conducted in this project for all eligible drivers.

Accident and violation data were compiled to evaluate treatment effectiveness over intervals up to 1 year following a "reference date" that was unique to each driver included in the pilot program. Chi-square tests compared the observed and expected frequencies of incidents for the treatment group, who received the SPE, and the comparison group, who received the BTW examination procedure previously administered by PennDOT.

The results of these analyses demonstrated clear and consistent benefits of the SPE, both in terms of violation and (chargeable) accident experience. At every month during the posttreatment evaluation period, a smaller percentage of drivers who passed the SPE had experienced either a single-vehicle or a multiple-vehicle accident than drivers who passed the BTW exam. This difference was statistically significant at the 3- and 6-month evaluation milestones for multiple-accident rates and was still marginally significant at 12 months ($p < .08$). Relative to the fraction of the comparison (BTW) group who had experienced multiple-vehicle accidents 1 year after treatment, the fraction of the treatment group with similar accident involvement was 20 percent lower. For all chargeable accidents, this difference was 16 percent.

An even more consistent and convincing reduction in all violations, and moving/point violations in particular, was found for the SPE. The lowered rates of convictions indicated by these data for the written examination versus the BTW exam were statistically significant at 3-, 6-, 9-, and 12-month milestones during the evaluation period.

These results reflect relative, not necessarily absolute, levels of effectiveness. A true (quasi-) experiment, with random assignment to treatment and control groups, was not a possibility in this research; furthermore, legal constraints in Pennsylvania ruled out the application of the SPE to a treatment group while a comparison group at the same (6-point) level of negligence received no intervention at all. Finally, an item analysis of the SPE to identify particular questions that were most predictive of subsequent driving behavior was not permitted, since hand scoring allowed only pass/fail status to be coded as an examination outcome.

At the same time, the observed decline in posttest accident and violation experience for passing drivers, coupled with the cost savings relative to administration of the prior, BTW examination procedure, was sufficiently encouraging to PennDOT

to support continued administration of the SPE in Pennsylvania, where current usage is projected at 50,000 or more drivers every year.

It was recommended by the report author that the SPE be automated to facilitate test administration, scoring, and record keeping. Presentation of the test items on a CRT could convey accident scenario information more clearly with fewer words, benefiting low-literacy and non-English-speaking drivers. Administrative costs would be lowered further, and security of the test forms would also be greatly enhanced. Additional research to determine whether a related examination approach with test content targeted at a specific user population—for example, young, inexperienced drivers—could demonstrate similar (relative) effectiveness also may be justified by the present findings.

ACKNOWLEDGMENTS

This research project was sponsored by PennDOT. Douglas Tobin, Rebecca Bickley, and Andrew Cleaver of PennDOT deserve special mention. In addition, the author gratefully acknowledges the contributions of Kathleen Knoebel of The Bionetics Corporation for software development, mainframe data extraction at PennDOT, and data analysis, and Kathy Lococo of The Scientex Corporation for her essential support in the project's conduct as well as the preparation of this manuscript.

REFERENCES

1. D. Meichenbaum. *Cognitive-Behavior Modification.* Plenum Press, New York, 1979.
2. L. Staplin, K. Lococo, and K. Knoebel. *Driver Improvement Index Pilot Study.* Report PA-91-021-85-05. Pennsylvania Department of Transportation, Harrisburg, 1992.

Publication of this paper sponsored by Committee on Operator Education and Regulation.

Evaluation of a National Traffic Police Force

DAVID M. ZAIDEL, IRIT HOCHERMAN, AND A. SHALOM HAKKERT

The evaluation of the National Traffic Police (NTP) in Israel, which was recently organized, is described. The creation of NTP entailed an organizational change in the structure of the traffic police force and a doubling of personnel and other resources. An independent, comprehensive evaluation was mandated to accompany the project. The evaluation program consists of four components: organizational analysis, monitoring of police operations, monitoring of traffic behavior, and analysis of accident trends. The methodology is described and results of the first phase of evaluation, which included the first three components, are presented. The major findings were as follows: inputs and outputs increased about twofold, roughly proportional to the increase in resources; no overall improvement in traffic behavior occurred. However, some places showed an improvement, and there was an indication of a relation between improvement in traffic behavior and the intensity of police presence. A higher-than-routine level of enforcement was necessary to achieve this improvement.

Police traffic operations in Israel used to be coordinated by local and district commanders of the national police force as part of their many other duties. At the district level there were personnel and subunits specializing in traffic-related services, but there was no national-level command, control, and communication structure dedicated to traffic enforcement and other traffic operations.

In September 1991 a National Traffic Police (NTP) force was established as an operational branch at the national police headquarters in Jerusalem, and all existing interurban traffic units came under its direct command. The transfer was gradual; the first phase, which is described in this paper, included two of the three rural districts, North and Center, covering about 2500 km of roads. These two districts include all the major, high-volume highways in Israel and account for about 80 percent of the injury accidents on interurban roads—2,600 injury accidents in 1991. The remaining South district was added in 1993.

NTP introduced an organizational change that put traffic police officers and traffic operations under a separate, nationally coordinated command. In addition, it entailed an increase in resources—manpower, vehicles, and associated enforcement equipment were nearly doubled in size.

NTP was empowered and expected to experiment with and modify strategies and tactics of traffic operations and active enforcement to achieve a higher level of traffic safety on Israeli roads. NTP was set up on a trial basis for 21 months; to our knowledge, it is one of the largest experiments in police traffic enforcement ever attempted. Most reported experiments or projects in this area tend either to be localized, general enforcement campaigns or to focus on specific target behaviors and populations, such as speeding or DWI offenders (*1–3*). Other experiments deal with particular variations in operational strategies, such as the impact of marked versus unmarked patrol cars (*4*).

In contrast, the NTP project was planned from the outset on a large geographic scale, comprehensive in scope, long range, and involving a permanent increase in enforcement resources rather than temporary shifts in their allocation. Given the national scope of the project, one might look not just for localized effects of enforcement, but also for apparent changes in the "normative standards of driving behavior" (*5*).

The Transportation Research Institute is evaluating the total operation of NTP to help improve NTP's functioning and to provide data for those required to determine its success. This paper describes the design and methodology of the evaluation and provides some results pertaining to the first 10 months of operation.

EVALUATION APPROACH, GOALS, AND OBJECTIVES

The underlying rationale of the evaluation plan is the relationship between traffic police activity and the level of traffic safety. It is hypothesized that an increase in resources and the organizational change will bring about an increase in enforcement, which, in turn, will increase both the objective and subjective probability of apprehension. Thus, deterrence and detection may cause changes in drivers' behaviors that are related to safety. These changes, in turn, will translate into a reduction in the number and severity of road accidents (*2,6*).

With this underlying assumption, NTP is viewed as an industrial plant working around the clock, using a variety of resources and producing interim and end products. The plant uses an "industrial process" of patrolling and enforcement to manufacture safety and efficient traffic flow. The ultimate indicator of the level of safety is the number and severity of accidents. Intermediate measures are behavioral patterns associated with safety, such as speed, conflicts at intersections, and seat belt use.

Evaluation is an integral part of modern industrial processes, which incorporate quality control and monitoring of inputs and outputs at all stages of production to optimize the production process and achieve the set goals and objectives. Similarly, the NTP evaluation plan assumes that the NTP structure and mode of operation are not rigidly defined, but rather evolve through a continuous feedback process. The

Transportation Research Institute, Technion—Israel Institute of Technology, Haifa 32000, Israel.

evaluation program, while not performed by NTP itself, is closely coordinated with this process and contributes to it.

In broad terms, the evaluation program was designed to monitor the organizational changes and their impact on NTP functioning and traffic operations, monitor changes in enforcement activity, measure changes in traffic or drivers' behavior and relate them to police actions, and, eventually, assess changes in traffic accidents that might be attributable to the new NTP. The evaluation program consisted of four components: organizational analysis, monitoring of NTP operations, monitoring of traffic behavior, and accident analysis.

This paper describes the first 10 months of NTP's operation. This period was deemed too short to induce discernible changes in accident trends. Therefore, accident analysis was deferred to the second phase of the evaluation.

ORGANIZATIONAL ANALYSIS

Methods

The organizational analysis entailed a structural and functional analysis of NTP at all levels of command, control, and field operation. The functions, units, stated goals and objectives, resources, procedures, tasks, and external and internal constraints of NTP were mapped and flowcharted according to various models of organizational behavior pertinent to industrial production and hierarchical organizations.

Various sources of information provided input to the organizational analysis. These include periodic visits to most NTP branches; interviews of key personnel in the NTP as well as in other branches of the police that were affected by it; and written records of command staff meetings, planning documents, and guidelines. A periodic questionnaire survey of NTP officers solicited their personal views on enforcement strategies and tactics, job satisfaction, the functioning of their units, and so forth. Data acquired during the monitoring of NTP operation, described in the next section, were also used to corroborate impressions and views obtained during site visits.

The organizational analysis was an important part of the first phase of the evaluation. It provided useful insights to the NTP and police command and contributed to organizational changes within NTP and the department of traffic. The major objectives of NTP were defined and prioritized, and operational measures of success were defined for each objective.

Findings

The analysis identified a number of bottlenecks in the process of enforcement within the police organization and inside related systems such as the judicial system. An example of a topic that was brought up during the organizational analysis is the process of active enforcement, or the production of citations. Data gathered and analyzed during the monitoring of police operations indicated that on the average, three to five citations are given by a patrol unit during a shift. Efforts to increase the level of enforcement brought the number up to seven citations per shift. The questionnaire revealed that the handling of a single citation requires a total of 20 to 30 min. This was corroborated during field visits. It was thus calculated that the limit to enforcement is set at a maximum of 10 citations per shift. This meant that a significant increase in enforcement cannot be achieved under the existing system, and the need for automatic enforcement was underlined. Further analysis of the production process of citations recorded by automatic equipment revealed bottlenecks down the line in units that were not accountable to NTP, such as the film processing lab and the central computer unit. It became clear that until all of these constraints are treated, no significant increase in enforcement is feasible. Thus, a complete revision of the processing of citations was suggested by the evaluation team and is now being examined by police officials.

The questionnaire surveys of police officers basically confirmed impressions that were formed during interviews and site visits, identified areas that acquired additional training, and demonstrated an overall high degree of job satisfaction and a positive attitude toward the new structure.

MONITORING OF TRAFFIC POLICE OPERATIONS

Methods

Monitoring of the NTP operations provided useful information for two aspects of the evaluation. First, it provided data on the efficiency of NTP traffic operations and the way they change with experience. Second, it provided quantitative measures of traffic operations, including active enforcement over time and across the road network, which could be linked with changes in drivers' behavior.

The main source of data was routine planning and reporting forms used regularly (on a daily or per shift basis) by the regional units. One such form, filled out daily, provides data on basic input parameters and efficiency measures, such as the number of active police officers and cars operating on each shift, the number of citations given, and the amount and types of enforcement equipment used and their output. Another form provided data on actual deployment of police cars on the road network for each shift. These forms were collected and coded regularly.

Another source of data was the files of the police central processing unit, which provided reliable data on citations given by NTP officers by type of citation, unit, time, and so forth.

Crude data on the number of NTP cars observed on the road were gathered by the evaluation team during trips to the traffic monitoring sites and during the observation there.

The data were analyzed to provide average monthly values of performance and efficiency indicators. These indicators included

- Input measures, such as number of patrol cars and police officers per shift;
- Efficiency measures, such as resource utilization rates for police officers, patrol cars, and enforcement equipment; and
- Intensity measures, such as total mileage driven by patrol cars, number of patrol cars observed per kilometer of road, rate of patrolling on each road section, total number of citations, and number of citations per patrol car per shift.

Findings

After a period of adjustment that lasted several months, all measures of activity showed an increase commensurate with the increase in resources. As of March 1992, activity measures reached a plateau at a level of about twice that of the pre-NTP period.

The main output measure is the number of citations given. This measure was analyzed by type of offense, by unit, area, time of day, and so forth. Figure 1 shows the total monthly citations given on the roads under NTP jurisdiction. An increase of about 100 percent was reached after a few months of operation. No evidence of any gain in efficiency is present—the increase in output can be attributed entirely to the increase in resources. Seasonal variation due to harsh weather (in the winter) and vacations (in the summer) is also apparent.

The routine deployment of patrol cars was based on a division of the road network into sections of 10 to 20 km each. Each section was assigned a priority level based on both traffic volumes and annual accident rates. A more detailed description can be found elsewhere (*7*). There are 118 such road sections under the jurisdiction of the two NTP regions, divided into two enforcement categories—A and B. Sections assigned Priority A should be patrolled every weekday shift, whereas sections in Category B should be patrolled two or three times a week. In addition, a special operations unit was formed in 1992 and provided extra coverage on a rotating basis to road sections that were considered especially unsafe.

After a period of adjustment, the deployment reached a level of 11 to 13 shifts per week on each Type A section. The level of coverage of Type B sections varied but was on the average less than two shifts per week. Sections that were reinforced by the special operations unit were patrolled by an average of at least 17 car shifts per week. On the average, a police car patrols the assigned section for 5 to 6 hr per 8-hr shift.

The independent observations made by the traffic monitoring team produced an estimate of four NTP patrol cars observed per 100 km of road (mostly Priority A roads).

Efficiency measures were mainly useful in illuminating problem areas in operation. For instance, the low rate of use of some of the enforcement equipment indicated the need for more training and precipitated some procedural changes in their use.

MONITORING OF TRAFFIC BEHAVIOR

Methods

The national scope of the NTP and the comprehensive nature of its operation required special methodological considerations with regard to what, where, when, whom, and how to observe and monitor. It was obvious that every desired driver behavior (or traffic characteristic) could not be monitored and that the monitoring could not be performed on every road and during all times. Yet the monitored behaviors should represent what happens on the road network much of the time and indicate whether there are changes in the behavior over time or locations. In addition, it has to be shown that such changes could be attributed to the NTP and that they are, eventually, related to improved safety.

The general approach in designing the monitoring system was to use a large sample of representative road sections and take repeated measures of the chosen behaviors. Some behaviors were automatically monitored over a number of hours, whereas more complex types of behaviors were manually collected over shorter periods, overlapping the autonomic records. Because of the importance of traffic volume in determining most other traffic characteristics, it was decided to always record traffic volumes along with other measurements.

The large number of observation sites recognized the large variability in traffic behavior due to local differences in traffic volume, vehicle mix, roadway characteristics, visibility, weather conditions, and other attributes. At the same time an effort was made to reduce the uncontrolled variance by limiting the variety of observation sites and following a procedure of repeated observations under as similar conditions as practically possible.

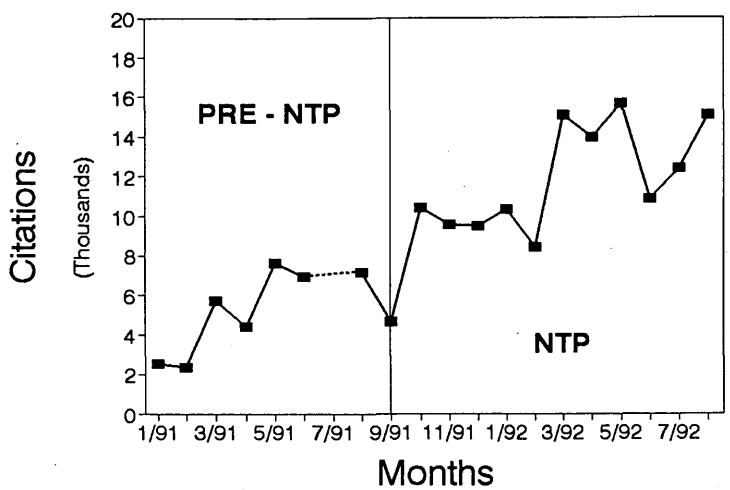


FIGURE 1 Number of citations produced each month.

Monitored Behaviors

The behaviors to be monitored were selected according to the following criteria:

- The behavior is enforced (at least in principle) by the police,
- The behavior is associated (at least in principle) with safety,
- The behavior can be reliably monitored and quantified, and
- The behavior can provide sufficient data for statistical analysis and inference.

The traffic behaviors that were selected for monitoring can be divided in two groups: (*a*) speed and following distance (measured on road sections) and (*b*) approach speed on the nonpriority road and the friction between turning vehicles, measured at nonsignalized T junctions.

Monitoring Sites

The sampling frame for selecting the observation sites was the list of 75 patrol sections with ADT greater than 5,000 vehicles. Twenty-four road segments were selected, 12 in each of the two NTP districts. The actual observation site in each road section was selected following an on-the-road survey. A nonsignalized junction with a sufficient volume of traffic was located and checked for the monitoring conditions. Traffic behavior at road sections was monitored on the main road leading to the junction but at a far enough distance to ensure independent behavior. The sampling, site selection, and actual observations were carried out independently of any planned or actual police activity at the road sections.

Monitoring Procedure

During Phase I, data were collected by the same team of observers, who visited each site five times, once every other month. Vehicle volumes, speeds, and following distances on the main road were recorded by electronic traffic counters. Counters were usually left for the whole day. Approach speed on the minor road was measured with a hand-operated portable radar (speed-gun) from inside a parked vehicle. Only free-moving vehicles were measured. At least 120 readings per measurement period were taken at each site.

The level of friction at the junction was assessed by an observer positioned at a convenient spot overlooking the junction. The observer counted and classified the turning maneuvers at the junction. Each turning maneuver was classified into one of three categories, depending on the amount of friction it generated in traffic at the time of its execution: normal, medium, or high. A minimum of 150 turning maneuvers per junction were observed at each measurement period.

Overview of Analysis

The purpose of the analysis was to detect meaningful changes in the behavioral measures of traffic, from the onset of the NTP project until the last (fifth) round of field measurements.

It took NTP a few months to restructure, reorganize, train new officers, acquire new vehicles, develop new operational procedures, and make other adjustments. Therefore, it was expected that the first round of observations would represent a baseline condition, after which some (hopefully positive) changes would take place as a result of NTP's stronger impact on traffic.

Specifically, the analyses were to find, across five measurement rounds, trends in the following summary measures:

- Average speed on main roads,
- Percentage of vehicles moving at speeds higher than the legal speed limit (90 km/hr),
- Percentage of following distances smaller than 1 sec or 2 sec,
- Average approach speed to a junction on the minor road, and
- Percentage of high and medium friction turning maneuvers at junctions.

As noted earlier, traffic volume has a major influence on the momentary values of other traffic characteristics. Therefore, automatic recordings of volumes and speeds were made for many hours, and the repeated data for each site were collected under conditions that were as similar as possible.

On the basis of data from all measurement rounds, a "common window of analysis" was identified for each site such that during similar hours of the day, during all measurement rounds, traffic volume was about the same. In most instances, it was a period of about 2.5 to 3.0 hr in the afternoon, just after the noon peak hour and before the next evening peak. The off-peak period ensured free-flowing traffic and allowed a better comparison of the behavior parameters over time.

A priori, there was only a very small chance for NTP to have a discernible impact on traffic behavior on a national scale. Therefore, the analysis was designed to enhance the possibility of finding any positive effects of enforcement on traffic behavior (if the effects were indeed there). The statistical analysis consisted of three steps. The first tested the hypothesis of a consistent change in behavior over time at a specific site. The second combined results from all sites to test the hypothesis of an overall change in behavior. These two steps were performed separately for each of the behavioral measures. The third step tested the association between police enforcement and changes in traffic behavior.

In the first step, statistical tests were performed at the single-site level. The occurrence of change in behavior was tested using three different comparisons within the five measurement periods: Period 1 versus Periods 2 through 5; Periods 1 and 2 versus Periods 3 through 5, and Periods 1 through 3 versus Periods 4 and 5. The choice of three different comparisons allowed the change in behavior to occur at any point in time after the first, baseline, measurement.

Mean speeds were tested using a one-way ANOVA with contrasts. The data of the other measures—percentage over speed limit, percentage following distances < 2 sec, percentage of high and medium friction at junction—were cross tabulated according to the different periods compared and tested with a χ^2 test for significance of the differences.

The results of the tests indicated, for each comparison, the direction of change: "positive" (a significant decrease in speed

or friction between vehicles, an increase in gaps), "negative" (a significant increase in speed, etc.) or "not significant."

The results of the three comparisons were combined into one measure of overall direction of change: positive (pos), negative (neg), or inconsistent (n.s). Table 1 gives an example of this procedure for mean speeds on the main road.

To test the hypothesis of an overall change in behavior, the results were combined over all sites. For each measure, the number of sites with positive change was compared with the number of sites with negative change, and a binomial test was used to test for a predominant trend.

To relate changes in traffic behavior to police enforcement, a crude measure of "enforcement level" was obtained for each site. It was based on the mean number of patrolling shifts per week assigned to the road section during the month preceding the field measurements. Each site was also assigned to one of three categories indicating an overall direction of change in traffic behavior. The measure of enforcement level was related to the overall trend of change at the site.

Findings

Changes in Speed

Table 1 is a summary table of mean speeds on the main road at 15 sites that are on two lane highways. The table also gives the trends of speed change on the basis of the ANOVA and tests for contrasts. The last column displays the overall trend. The binomial test for the significance of having 7 out of 10 comparisons positive by chance shows that it is likely at a probability of $p = .17$ (one-tail). Examination of the table indicates that the differences between measurement rounds are usually small—less than 5 percent.

Approach speeds on the minor roads have changed even less than speeds on the major roads connecting with them. For 21 sites with relevant data, 9 showed a decrease in speed over time, 5 showed an increase, and for 7 there was no clear change. There is clearly no evidence of a predominant trend of change here.

The percentage of vehicles going over 90 km/hr (speed limits are 80 km/hr) varied from as low as 1 to as high as 35 percent over the 15 two-lane sites. In all but five of the sites there was no significant trend of change. The expectation that this measure would be more sensitive to enforcement effect was not fulfilled. In retrospect, it is not surprising in view of the inevitable larger variance in the high tail end of a speed distribution. The mean is actually a more reliable measure.

Changes in Percentage Following Too Closely

Generally, there was no significant change in the values of these measures across the five periods. Only a few of the individual comparisons were significant in the χ^2 test and, consequently, the predominant trend was "no change."

The percentage of vehicles following at a distance of 2 sec or less hovered around the 20 percent value. It was clearly volume dependent: up to 27 percent at the site with the highest traffic volumes and down to 10 percent at the lower volume scale.

Changes in Level of Friction at Junctions

The friction data were aggregated across all turning maneuvers. The percent of medium- and high-level encounters ranged from 1 to 10 percent. At 17 out of 25 junctions there was no significant change between periods. Of the eight sites that showed a relatively consistent trend of change, seven were in a positive direction, less friction at later periods of measurement (one-tailed $p = .035$.)

TABLE 1 Mean Speed by Site, Measurement Period, and Direction of Speed Change

Site ID	Mean speed (km/h) on main road					Direction of speed change			
	Period 1	Period 2	Period 3	Period 4	Period 5	p1/p2-p5	p1-2/p3-5	p1-3/p4-5	Overall
1	74.85	76.33	76.33	75.36		n.s	n.s	neg	neg
3	79.71	72.99	77.97	80.42	76.85	neg	neg	pos	n.s
4	75.75	74.22	73.69	74.00		pos	pos	pos	pos
5	75.29	71.82	70.26	72.88		n.s	pos	pos	pos
6	78.13	77.79	77.53	76.27	78.78	n.s	n.s	n.s	n.s
7	75.11	76.12	74.23	74.21	74.40	pos	pos	n.s	pos
8	67.08	68.08	65.42	65.17	64.07	pos	pos	pos	pos
9	66.73	67.59	67.49	66.79	68.59	n.s	n.s	n.s	n.s
10	79.58	76.30	77.22	76.56		pos	pos	pos	pos
11	70.50	73.07	66.39	70.89	63.51	pos	pos	pos	pos
12	83.06	79.95	82.36	81.61		n.s	n.s	pos	n.s
15	67.98	73.41	67.74	68.08	69.37	n.s	pos	neg	n,s
16	70.45	70.30	69.05		71.83	neg	n.s	n.s	neg
17	83.65	83.67	86.38	82.75	83.72	pos	n.s	n.s	pos
23	66.22	65.61	65.97	66.15	67.80	neg	n.s	n.s	neg

Did Enforcement Influence Traffic Behavior?

So far, analysis of the results for each measure and across all relevant sites showed no significant changes between periods, with the possible exception of the friction measure. However, it was also shown that there were large differences between sites. Some showed a fairly consistent trend of improvement on all or some of the measures, many showed no change or an inconsistent trend, and a few showed a negative trend of change. Is there some common feature between the subgroups of sites that can explain the differences? Table 2 presents an attempt to relate the predominant trend of change at a site to the level of enforcement.

The 20 sites with data on most of the measures for most of the periods were classified according to the trend of change on all the measures into three categories: positive, no change, and negative. Sites with at least two behaviors that displayed a positive change and no behavior that displayed a negative change were defined as sites with an overall improvement in traffic behavior. Sites with at least two negative changes and no positive change were assigned to the category of negative change. All other sites were assigned to the group of no consistent change.

The same sites were independently classified on the basis of police operational records according to the level of enforcement that was assigned to the road section that included the site. The level of enforcement was measured by the mean number of patrolling shifts (typically a single patrol car for a 10- to 20-km section) per week. The medium level of enforcement, 11 to 13 shifts per week, represented the standard enforcement practiced by the police. Lower-priority sections were patrolled less frequently. The high level of enforcement represented a special concentrated effort on selected routes for a couple of months before the field measurements.

Table 2 indicates the number of sites associated with each level of enforcement and the direction of change that has occurred in the behavioral measures. The sites on roads that were subjected to a relatively high level of enforcement tended to show mainly a positive effect (five out of seven), sites with the medium, standard enforcement concentrated in the no change category (also five out of seven), and the six sites with an enforcement level lower than standard split evenly between the three categories of change.

TABLE 2 Sites, by Level of Enforcement and Direction of Change in Traffic Behavior

Level of Enforcement	Positive change	No change	Negative change	Total
HIGH 14+	5	1	1	7
MEDIUM 11-13	0	5	2	7
LOW > 10	2	2	2	6
TOTAL	7	8	5	20

The result is intriguing but must be considered very tentative in view of the small number of cases in the table. The relationship suggests that it may take much more than a low or medium level of enforcement to have a reasonable chance of influencing traffic behavior in terms of the measures evaluated in this study. There is some indication in the data that concentrated police activity in a generally targeted area may improve traffic behavior.

Since it is not possible with limited resources to boost the level of police activity everywhere, all of the time, it is also clear that more attention must be given to what police are doing in the process of enforcement and how to increase its impact apart from increasing manpower and vehicles on each road.

ACKNOWLEDGMENT

This study was cosponsored by the Highway Safety Administration of the Ministry of Transport, the Police, and the Treasury Department.

REFERENCES

1. M. Armour. The Effect of Police Presence on Urban Driving Speeds. *Australian Road Research*, Vol. 14, No. 3, 1984.
2. G. Nilsson. Speed Limits, Enforcement and Other Factors Influencing Speed. In *Enforcement and Rewarding Strategies and Effects* (M. J. Koornstra and J. Christensen, eds.). Leidschendam, the Netherlands, 1991.
3. D. M. Zaidel. A Modelling Perspective on the Culture of Driving. *Accident Analysis and Prevention*, Vol. 24, No. 6, 1992.
4. D. Shinar and A. J. McKnight. *The Effects of Enforcement and Public Information on Compliance.* New York, 1985.
5. J. A. Rothengatter. Normative Behaviour Is Unattractive If It Is Abnormal: Relationships Between Norms, Attitudes and Traffic Law. In *Enforcement and Rewarding Strategies and Effects* (M. J. Koornstra and J. Christensen, eds.). Leidschendam, the Netherlands, 1991.
6. T. Makinen, L. Beilinson, and M. Salusjarvi. Traffic Enforcement Strategies and Tactics. Presented at 2nd International Conference on New Ways and Means for Improved Road Safety and Quality of Life, Tel Aviv, Israel, Oct. 1991.
7. A. S. Hakkert, A. Yelinek, and E. Efrat. Police Surveillance Methods and Police Resource Allocation Models. In *Enforcement and Rewarding Strategies and Effects* (M. J. Koornstra and J. Christensen, eds.). Leidschendam, the Netherlands, 1991.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Statistical Methods To Support Induced Exposure Analyses of Traffic Accident Data

GARY A. DAVIS AND YIHONG GAO

When it is possible to identify the drivers involved in two-vehicle accidents as either at fault or innocent, induced exposure methods offer a way to assess the relative accident risk of driver subgroups, even when group-specific measures of exposure are unavailable. A cross tabulation of two-vehicle accidents by group membership of the at-fault and victim drivers forms a contingency table, and statistical methods derived from contingency table analysis can be used to make inferences concerning the variables arising in the induced exposure model. It is shown how the standard contingency table test for independence of row and column classifications provides a test of the assumption that the victims are sampled randomly and how an odds ratio statistic can be used to estimate the ratio of the accident rates between two driver subgroups. This estimator is asymptotically normally distributed, and a formula is given for estimating its standard error. An Empirical Bayes method for identifying sites where one driver subgroup has a significantly higher accident rate than does another is then presented. These procedures are illustrated using several actual accident data sets.

Over the past several years, the traffic safety community has shown an increased interest in assessing the accident risk of particular driver subgroups. The main emphasis has been on older drivers (*1*), but recently attention has also been given to younger drivers (*2*) and to mounting evidence for an increasing number of accidents involving women drivers (*3*). Unfortunately, the study of such problems is made difficult by the fact that an increase in accidents for some driver subgroup can be attributed either to an increase in the tendency of that group to have accidents (its accident rate), to an increase in that group's opportunity to be involved in accidents (its exposure), or to some interaction between these factors. Using accident counts to make inferences concerning a subgroup's accident rate will generally require knowledge of that group's exposure, but measures of exposure are difficult to define in a completely satisfactory way (*2*) and are even more difficult to obtain in disaggregated forms. For instance, even if we can agree that a variable, such as vehicle kilometers of travel (VKT), is the appropriate exposure measure for older drivers, estimating the VKT for this subgroup usually requires asking a sample of drivers to estimate the number of kilometers they have driven during the past year. And even when such data are available, they usually tell us little about the exposure of a subgroup on smaller areal units, such as a highway corridor or a single intersection.

Safety researchers have been aware of these difficulties for at least 25 years, and in the early 1970s induced exposure methods were presented as providing at least a partial solution to this problem (*4*). Following some intense initial interest, induced exposure methods appear to have suffered a period of neglect, but recently several papers have used these ideas to investigate accident risk to older drivers (*5–8*). The induced exposure model assumes that in a majority of two-vehicles accidents, one driver can be identified as the at-fault driver, whereas the other is treated as an innocent victim. Innocent victims are assumed to be "selected" by the at-fault driver randomly from the pool of available drivers, with the probability that the innocent victim is the member of a given subgroup being directly proportional to that subgroup's exposure at the accident site. Thus the same measure of exposure reflects both a subgroup's opportunity to cause accidents and its opportunity to be involved as victims. From comparisons of the proportion of accidents that a subgroup causes with the proportion in which it is involved as innocent victims, it is possible to identify subgroups that have accident rates higher or lower than the average for all groups (*9*).

But before the promise of induced exposure methods can be fully realized, it is necessary to answer several questions related to the statistical properties of induced exposure measures. First, the assumption of random selection of victims appears somewhat controversial (*10*), although studies investigating its validity have tended to support it (*5,10*). Still, it would be useful if a test of the tenability of this assumption could be conducted on any given data set. Second, recent studies have used the induced exposure method in essentially a deterministic manner, treating data-dependent quantities as if they were known with certainty, with no attempts made to estimate likely ranges of error. When one has very large data sets, it may be possible to invoke the law of large numbers to justify ignoring random effects, but many, if not most, applications of induced exposure will likely involve more modest data sets. Here it would be useful to have procedures for determining confidence bounds and testing hypotheses for induced exposure estimates. Finally, many safety engineers are ultimately responsible for deciding on particular safety improvements for particular locations. If these improvements are targeted at a specific driver subgroup, it may be necessary to identify specific locations where that subgroup is at heightened risk. Aggregated data will not generally provide this

Department of Civil and Mineral Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, Minn. 55455.

level of detail, necessitating an extension of induced exposure ideas to the problem of identifying high-hazard locations.

In what follows, we will first show a natural correspondence between statistical inference using induced exposure ideas and more standard methods of contingency table analysis. This will lead first to a straightforward test as to whether the assumption of random selection of victims is tenable for a given data set and then to a method for computing maximum likelihood estimates of the ratio of the accident rates for two driver subgroups. We point out that the natural logarithm of this estimated rate-ratio is approximately normally distributed and give a formula for estimating its standard error. These methods are then illustrated using two actual traffic data sets. We next turn to the problem of identifying high-hazard locations and derive a Bayes estimator for the log rate-ratio, along with its posterior standard error. Given data from a number of sites, we then show how an Empirical Bayes approach can be used to compute point estimates and approximate confidence intervals for the log rate-ratios for each site. The paper ends with conclusions and recommendations for further research.

INDUCED EXPOSURE AND CONTINGENCY TABLES

We begin with a more formal statement of the induced exposure model. Suppose the driver population has been divided into m subgroups, and let n_i denote the number of accidents involving Driver Subgroup i in some area (such as a city) over some time interval (such as a year). Assume n_i to be the outcome of a Poisson random variable, with mean $\lambda_i E_i$, where λ_i is the accident rate for Driver Subgroup i and E_i is the exposure for Driver Subgroup i. If the exposure values E_i are known exactly, the maximum likelihood estimates of the accident rates λ_i are given by

$$\hat{\lambda}_i = \frac{n_i}{E_i}$$

and assessment of the relative risk to driver subgroups can be based on these estimated accident rates. But as noted earlier, group-specific measures of exposure are difficult to estimate reliably. To implement an induced exposure approach, it is first assumed that in a majority of two-vehicle accidents, one driver can be considered to have caused the accident, whereas the other is assumed to be an innocent victim. The at-fault drivers are assumed to cause accidents according to the Poisson accident model, whereas the subgroup of the victim is assumed to be selected randomly, with probability of selection being directly proportional to the group exposures. Defining $r_i = E_i/\Sigma_k E_k$, the probability the victim is chosen from Subgroup i, and n_{ij} = number of accidents for which the at-fault driver came from Subgroup i while the victim came from Subgroup j, it follows that the n_{ij} are the outcomes of independent Poisson random variables with mean values $r_j \lambda_i E_i$. By taking the total number of accidents in the sample, $n = \Sigma_i \Sigma_j n_{ij}$, as fixed and defining $p_i = \lambda_i E_i / \Sigma_k \lambda_k E_k$, it can be shown that the n_{ij} are now the outcomes of a multinomial random vector with number of "trials" equal to n and the probability of a given two-vehicle accident having an at-fault driver from Subgroup i and a victim from Subgroup j being simply $p_i r_j$. The n_{ij} can be thought of as entries in a cross-tabulation table, where two-vehicle accidents are classified according to the group membership of the at-fault and victim drivers. For example, Table 1 gives the expected cell counts (the expected values of the n_{ij}) and marginal probabilities for the case where only two subgroups, denoted by 1 and 2, are of interest.

In Table 1 the probability that a given two-vehicle accident falls in a cell is simply the product of the corresponding row and column marginal probabilities, so that the table shows statistical independence between its row and column classifications. This structure is a consequence of the assumption that the subgroup of the victim is selected randomly, and the standard tests of independence provide methods for identifying data sets for which this assumption is not valid. As an example, for a 2 × 2 table such as that given in Table 1, it is well known (*11*) that under the hypothesis of independence, the log cross-product ratio statistic

$$\hat{\theta} = \log_e\left(\frac{n_{11}n_{22}}{n_{12}n_{21}}\right)$$

has, for large values of the sample size n, approximately a normal distribution with a mean of zero and a variance that can be estimated by

$$\hat{\sigma}_\theta^2 = \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}$$

Tests for random selection of victims can then be conducted using the standard normal, or z, distribution.

Assuming now that the data in an induced-exposure table satisfy the assumption of random selection of victims, we turn to the problem of making inferences concerning the accident

TABLE 1 Example 2 × 2 Induced Exposure Table

		Innocent Victim		
		1	2	
	1	$E[n_{11}]=np_1r_1$	$E[n_{12}]=np_1(1-r_1)$	p_1
At Fault				
Driver	2	$E[n_{21}]=n(1-p_1)r_1$	$E[n_{22}]=n(1-p_1)(1-r_1)$	$1-p_1$
		r_1	$1-r_1$	

rates λ_i. To simplify some of the following notation, we define the marginal totals

$$x_i = \sum_k n_{ik}$$

$$y_j = \sum_k n_{kj}$$

A straightforward application of maximum likelihood methods yields the ML estimators of p_i and r_j,

$$\hat{p}_i = \frac{x_i}{n}$$

$$\hat{r}_j = \frac{y_j}{n}$$

Unfortunately, since $\Sigma_i p_i = \Sigma_j r_j = 1$, the induced exposure table is completely characterized by $2(m - 1) + 1$ parameters, making it impossible to uniquely identify the $2m$ parameters λ_i and E_j. It is possible, however, to estimate and compare relative quantities, p_i for example being the ratio of the expected number of accidents caused by Subgroup i to the total expected number of accidents, whereas r_j is the ratio of the exposure for Subgroup j to the total exposure. The measure that has appeared most often in the literature is the involvement ratio (*5–10*)

$$\mathrm{IR}_i = p_i/r_i$$

with $\mathrm{IR}_i = 1.0$ being taken as evidence that the accident rate for Group i is typical of the whole population. This interpretation follows by noting that the accident rates should be independent of the exposures, so that $\mathrm{IR}_i = 1.0$ and $E_j = E$ for each j implies

$$\lambda_i = (1/m) \sum_j \lambda_j$$

When only two subgroups are available (i.e., $m = 2$), $\mathrm{IR}_1 = 1.0$ is equivalent to $\lambda_1 = \lambda_2$. Alternatively, as elsewhere (*12*), one could consider the difference $p_i - r_i$, with $p_i - r_i = 0$ having the same interpretation as $\mathrm{IR}_i = 1.0$.

The involvement ratio allows the analyst to identify which subgroups have accident rates that exceed the populationwide average but does not provide readily interpretable information concerning the magnitude of this discrepancy, nor does it provide a means for comparing the relative accident risks of two different subgroups. However, the ratio of two involvement ratios has the form of an odds ratio statistic and is equal to the ratio of the respective accident rates:

$$\frac{\mathrm{IR}_i}{\mathrm{IR}_j} = \frac{p_i r_j}{p_j r_i} = \frac{\lambda_i}{\lambda_j}$$

If we define the log rate-ratio statistic as

$$\Delta_{ij} = \log_e\left(\frac{\lambda_i}{\lambda_j}\right)$$

it is straightforward to verify that the ML estimate of Δ_{ij} can be computed via

$$\hat{\Delta}_{ij} = \log_e\left(\frac{x_i y_j}{x_j y_i}\right)$$

and an application of the delta method yields that, for large n, the distribution of Δ_{ij} is approximately normal, with mean equal to $\hat{\Delta}_{ij}$ and variance that can be estimated via

$$\hat{\sigma}^2_\Delta = \frac{1}{x_i} + \frac{1}{y_j} + \frac{1}{x_j} + \frac{1}{y_i}$$

This last result is a consequence of the fact that, given row and column independence, the likelihood function of the induced exposure table factors into two components, one being proportional to the marginal likelihood of the row totals and the other being proportional to the marginal likelihood of the column totals. This provides a method for testing hypotheses concerning Δ_{ij} and for constructing approximate confidence intervals for the rate-ratio λ_i/λ_j.

As an illustration of the utility of these methods, first consider the data given in Table 2, originally presented by Lyles et al. (*10*). Here we have two 2×2 induced-exposure tables, with the driver subgroups being male and female. The upper table gives the cross tabulation for non–rush hour daytime interstate accidents in Michigan for 1988, and the lower table is a similar cross tabulation of nighttime interstate accidents. Testing first whether the assumption of random victim selection is tenable for these tables (i.e., that the estimated log cross product ratio, $\hat{\theta}$, is not significantly different from zero), we obtain for the upper table $\hat{\theta} = -0.039$, $z = -0.502$, $p > .6$, whereas for the lower table we obtain $\hat{\theta} = 0.055$, $z = 0.65$, $p > .5$. In both cases, independence of row and column classifications appears tenable. Next, we consider whether the accident rate for males is greater than that for females by testing the null hypothesis $\lambda_m = \lambda_f$ against the one-sided al-

TABLE 2 Example Induced Exposure Tables from Lyles et al. (*10*)

		Day Time Non-rush Hour	
		Innocent Victim	
		Male	Female
	Male	1810	941
At Fault			
Driver	Female	678	339
		Night Time	
		Innocent Victim	
		Male	Female
	Male	2232	894
At Fault			
Driver	Female	605	256

ternative $\lambda_m > \lambda_f$. The log rate-ratio provides an appropriate test statistic, and for the upper table we obtain $\hat{\Delta}_{mf} = 0.33$, $z = 6.57$, $p < .001$, and for the lower table we obtain $\hat{\Delta}_{mf} = 0.386$, $z = 7.43$, $p < .001$, indicating that, in both cases, the accident rate for males is significantly higher than that for females. For the upper table, an approximate 90 percent confidence interval for the rate-ratio λ_m/λ_f would be (1.28, 1.51), whereas a similar confidence interval for the lower table would be (1.35, 1.60). For both tables, it appears that the accident rate for male drivers is around 40 percent higher than that for female drivers.

As a second example, consider the data presented in Table 3. Here, drivers are divided into two subgroups according to age, with Group 1 being middle-aged drivers (ages 25 to 55) and Group 2 being "older" drivers (ages 56 and over). The upper table presents a cross tabulation of two-vehicle accidents occurring at the signalized intersections along a section of Minnesota Trunk Highway (MNTH) 47 during 1988–1989. The lower table presents a similar cross tabulation for MNTH 65, which runs about 1.5 km east and parallel to MNTH 47. Checking first to see whether the assumption of random victim selection is tenable for these two tables, we obtain for MNTH 47 $\hat{\theta} = -0.419$, $z = -0.93$, $p > .34$. For MNTH 65 we obtain $\hat{\theta} = -0.378$, $z = -1.08$, $p > .28$. Again, the random selection assumption appears acceptable. Testing next for whether older drivers have higher accident rates than do middle-aged drivers, we obtain for MNTH 47 $\hat{\Delta} = 0.2$, $z = .84$, $p > .20$, and for MNTH 65 we obtain $\hat{\Delta} = 0.28$, $z = 1.50$, $p < .07$. Thus the data from MNTH 47 show no evidence for increased accident risk to older drivers, but the data from MNTH 65 give a somewhat tentative suggestion that older drivers have higher accident rates. This sort of information could be useful to a safety engineer responsible for programming safety improvements.

TABLE 3 Induced Exposure Tables from Two Minnesota Highways

		MNTH 47	
		Innocent Victim	
		Middle-Aged	Older
	Middle-Aged	131	34
At Fault Driver	Older	41	7

		MNTH 65	
		Innocent Victim	
		Middle-Aged	Older
	Middle-Aged	202	52
At Fault Driver	Older	68	12

EMPIRICAL BAYES IDENTIFICATION OF HIGH-HAZARD LOCATIONS

The second example presented above suggested that the MNTH 65 corridor might be a candidate for safety improvements targeted at older drivers. But since there are 29 signalized intersections providing data for that example, it could very well be that these sites differ in the risk they pose to older drivers. Because the numbers of accidents occurring at particular sites over a 2- or 3-year period typically tend to be in the range 0 to 50, the uncertainty attached to site-specific ML estimates tends to be high, and application of the asymptotic statistical methods described earlier to individual sites is questionable. Alternatively, identifying high-hazard locations can be viewed as an example of a multiparameter estimation problem, so that Empirical Bayesian (EB) statistical methods might profitably be employed (*13,14*); in fact Davis and Koutsoukos (*12*) have described an EB approach for estimating the difference $p_i - r_i$. Here, we describe how EB estimates and confidence intervals can be computed for the log rate-ratio statistic defined above. To simplify the presentation of some of the following equations, we will restrict our attention to the case in which only two driver subgroups are of interest.

Let the two driver subgroups of interest be denoted by 1 and 2 and assume that there is available a 2 × 2 induced-exposure table for each of a set of N sites making up our sample. Let the individual sites be indexed by $k = 1, \ldots, N$, and define the variables

p_k = probability that an accident at Site k had a driver from Subgroup 1 as the at-fault driver,
r_k = probability that an accident at Site k had a driver from Subgroup 1 as the innocent victim,
n_k = total two-vehicle accidents available for Site k,
x_k = number of accidents from Site k where the at-fault driver was from Subgroup 1, and
y_k = number of accidents from Site k where the innocent victim was from Subgroup 1.

The EB model assumes that the actual accident counts for a site are generated by a two-stage random process. First, the probabilities p_k are randomly assigned to sites as the outcomes of independent, identically distributed (iid) Beta random variables, with means and variances given by

$$E[p_k] = p$$

and

$$Var[p_k] = p(1 - p)/(m_1 + 1)$$

The r_k are assigned as iid Beta random variables with means and variances

$$E[r_k] = r$$

and

$$Var[r_k] = r(1 - r)/(m_2 + 1).$$

Given p_k, r_k, and n_k, the accidents are then assigned to cells in the induced exposure table according to the multinomial

model described earlier. The log rate-ratio statistic for Site k becomes

$$\Delta_k = \log_e\left[\frac{p_k(1 - r_k)}{r_k(1 - p_k)}\right]$$

In a manner similar to that used by Maritz (*15*), it can be shown that if the underlying prior parameters p, m_1, r, and m_2 are known in advance, the posterior means and variances of the Δ_k are given by

$$E[\Delta_k|n_k, x_k, y_k, m_1, p, m_2, r] = \Psi(m_1p + x_k)$$
$$+ \Psi[m_2(1 - r) + n_k - y_k] - \Psi[m_1(1 - p) + n_k - x_k]$$
$$- \Psi(m_2r + y_k)$$
$$Var[\Delta_k|n_k, x_k, y_k, m_1, p, m_2, r] = \Psi'(m_1p + x_k)$$
$$+ \Psi'[m_1(1 - p) + n_k - x_k]$$
$$+ \Psi'(m_2r + y_k) + \Psi'[m_2(1 - r) + n_k - y_k] \quad (1)$$

where $\Psi(x)$ denotes the digamma function and $\Psi'(x)$ denotes the trigamma function:

$$\Psi(x) = \frac{d\log_e[\Gamma(x)]}{dx}$$
$$\Psi'(x) = \frac{d\Psi(x)}{dx} \quad (2)$$

The expression for the posterior variance of Δ_k follows from the fact that the joint posterior distribution of p_k and r_k factors into two components, one containing m_1, p, and x_k and the other containing m_2, r, and y_k. When numerical software for evaluating these functions is not available, they can be approximated using the first-order terms of their asymptotic expansions (*16*)

$$\Psi(x) \approx \log_e(x) - \frac{1}{2x}$$
$$\Psi'(x) \approx \frac{1}{x} \quad (3)$$

Furthermore, the posterior distribution of the Δ_k is well approximated by a normal distribution with means and variances given in Equation 1, so that if the prior parameters m_1, p, m_2, and r are known, point and interval estimates of the Δ_k can be computed using either Equation 1 or Equation 3.

In practice, though, the prior parameters p, m_1, r, and m_2 will not be known and must also be estimated from data. The EB approach proceeds by simply replacing the prior parameters in Equation 1 with these estimates, so that the EB estimate of Δ_k is

$$\hat{\Delta}_k = \Psi(\hat{m}_1\hat{p} + x_k) + \Psi[\hat{m}_2(1 - \hat{r}) + n_k - y_k]$$
$$- \Psi[\hat{m}_1(1 - \hat{p}) + n_k - x_k] - \Psi(\hat{m}_2\hat{r} + y_k) \quad (4)$$

and the EB estimate of the variance of Δ_k is

$$\hat{\sigma}_k^2 = \Psi'(\hat{m}_1\hat{p} + x_k) + \Psi'[\hat{m}_1(1 - \hat{p}) + n_k - x_k]$$
$$+ \Psi'(\hat{m}_2\hat{r} + y_k) + \Psi'[\hat{m}_2(1 - \hat{r}) + n_k - y_k] \quad (5)$$

An EB confidence interval with approximate coverage probability $1 - \alpha$ would then be $(\hat{\Delta}_k - z_{\alpha/2}\,\hat{\sigma}_k, \hat{\Delta}_k + z_{\alpha/2}\,\hat{\sigma}_k)$.

Maximum likelihood estimates of the parameters p, m_1, r, and m_2 can be found as values maximizing the marginal distribution

$$L(p, m_1, r, m_2) = \prod_{k=1}^{N}\left\{\frac{n_k!}{\prod_{ij} n_{ij,k}!}\frac{B[m_1p + x_k, m_1(1 - p) + n_k - x_k]}{B[m_1p, m_1(1 - p)]}\right.$$
$$\left.\times\frac{B[m_2r + y_k, m_2(1 - r) + n_k - y_k]}{B[m_2r, m_2(1 - r)]}\right\} \quad (6)$$

Here $B(a, b)$ denotes the Beta integral evaluated at a and b. Computation of the estimates is simplified by the fact that Equation 6 factors into two components, one containing p and m_1 and the other containing r and m_2, so that the maximization problem decomposes into two bivariate problems.

One problem that can arise in practice is that the likelihood function (Equation 6) may be unbounded with respect to either m_1 or m_2 (i.e., no finite MLE may exist for these parameters). This was in fact the case for the parameter m_1 for both the MNTH 65 and the MNTH 47 data sets. The simplest solution to this problem is to constrain the parameters m_1 and m_2 to be less than some appropriately large value and use this bound as the MLE in those situations where the MLE is unbounded. To arrive at a plausible upper bound, recall that the objective of this method is to identify locations where the accident rates for Groups 1 and 2 satisfy $\lambda_1 > \lambda_2$. Using the formulas in Equation 1 coupled with the normal approximation of the posterior distribution of Δ_k, it is possible to express the posterior probability that $\lambda_1 > \lambda_2$ as a function of m_1 and m_2. By inserting the MLE for m_2 into this function and then plotting this probability as a function of m_1, it is possible to gain an idea of the sensitivity of the final decision to the choice of an upper bound. Figure 1 shows such plots for four typical intersections selected from MNTH 65. In each

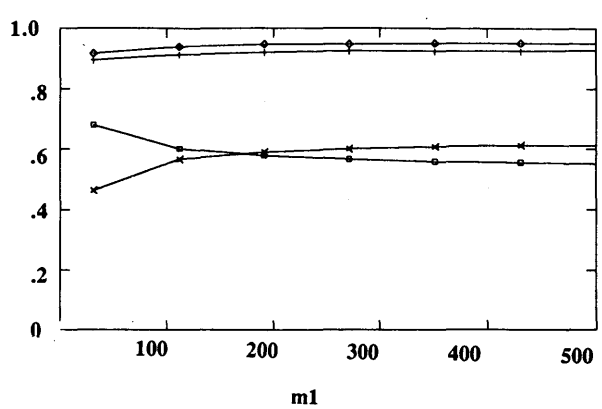


FIGURE 1 Approximate posterior probability that $\lambda_1 > \lambda_2$ as a function of m_1, for four intersections on MNTH 65.

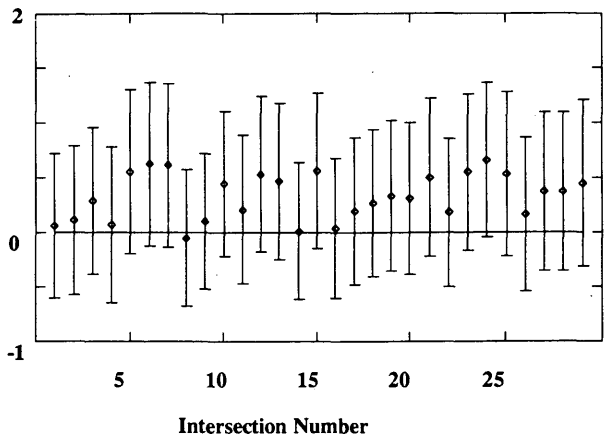


FIGURE 2 EB point estimates of log rate-ratios for 29 intersections on MNTH 65, along with approximate 90 percent EB confidence intervals. Upper bound for m_1 set at 100.

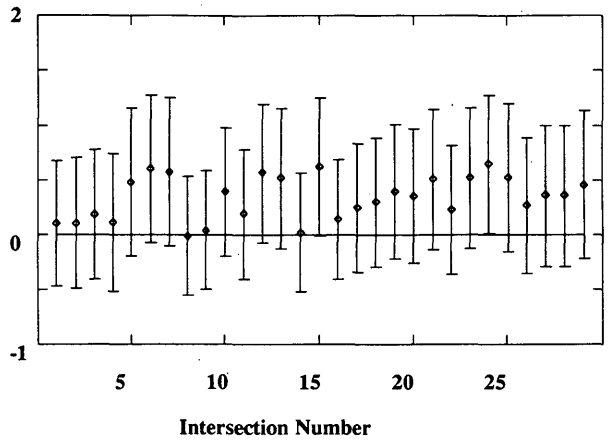


FIGURE 4 EB point estimates of log rate-ratios for 29 intersections on MNTH 65, along with approximate 90 percent EB confidence intervals. Upper bound for m_1 set at 500.

of these cases, when $m_1 > 200$ the posterior probability tends to stabilize into a slowly monotonic function of m_1. This pattern was present in each of the sites included in this study, with most of the change in posterior probability tending to occur for m_1 less than 500 and values of m_1 beyond 500 tending to produce fairly small changes in posterior probability.

To illustrate this EB approach, we return to the MNTH 65 example presented earlier. There were a total of 29 signalized intersections along this section of MNTH 65, and the induced exposure data presented in Table 3 were disaggregated according to the intersection where the accidents took place. Two MATHCAD 3.0 computational documents were developed. The first computed bounded ML estimates of p, m_1, r, and m_2 via Equation 6, with upper bounds being user-specified inputs, and then wrote these estimates to a file. The second document read these estimates, computed the EB estimates for Δ_k, σ_k, and approximate 90 percent confidence intervals for Δ_k, for each of the 29 intersections, and then created the graphs shown in Figures 2 through 4. The upper bound for

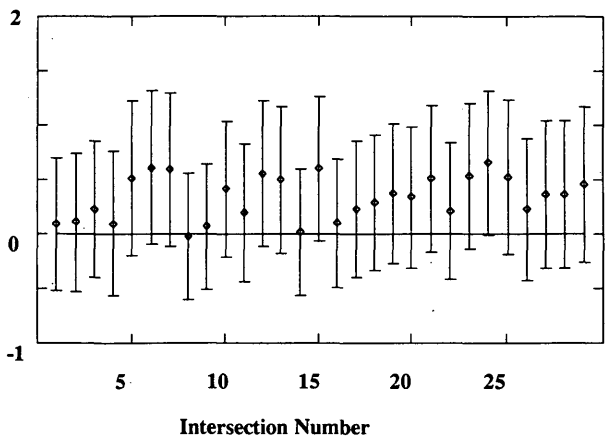


FIGURE 3 EB point estimates of log rate-ratios for 29 intersections on MNTH 65, along with approximate 90 percent EB confidence intervals. Upper bound for m_1 set at 200.

m_1 was set at 100 in Figure 2, at 200 in Figure 3, and at 500 in Figure 4. Since an estimated $\hat{\Delta}_k$ that is not significantly different from zero indicates a site where the accident rate for older drivers is not significantly different from that of middle-aged drivers, inspection of Figures 2 through 4 indicates that the risk to older drivers is not evenly distributed along the roadway. If present at all, it appears concentrated on two segments, one containing Intersections 5, 6, and 7 and the other containing Intersections 23, 24, and 25. This qualitative identification appears to be robust with respect to the upper bounds placed on m_1. These results are similar to those presented for the same data set, but a somewhat different computational method, elsewhere (*12*).

CONCLUSION

In this paper we have formalized some of the relationships between induced exposure and contingency table analyses, used these results to identify a test for the random selection of accident victims, and then developed an estimator for the ratio between the accident rates for two different driver subgroups. An Empirical Bayes approach was then presented for estimating these rate-ratios for each of a number of accident sites and using approximate confidence intervals around these estimates to identify locations where a given driver subgroup might be at increased risk. The utility of these procedures was illustrated using actual traffic accident data.

Although certainly not a panacea, the induced exposure model offers a promising approach for estimating the differential in accident risk experienced by subgroups of drivers, and it is hoped that the statistical methods described here will facilitate a wider use of and research into induced exposure methods. Of particular interest would be an extension of this approach to multiway cross-tabulation tables, permitting the analyst to assess the effect of possible causal factors on accident rate differentials. A special case would be the problem of assessing the impact of safety countermeasures, using before and after data. Finally, user-friendly implementations of these methods are probably needed to facilitate their widespread adoption.

ACKNOWLEDGMENTS

The authors would like to thank Susan Scharenbroich and Dave Miller of the Minnesota Department of Transportation (MNDOT) for their assistance in conducting this research. This project was supported by MNDOT.

REFERENCES

1. *Special Report 218: Transportation in an Aging Society: Improving Mobility and Safety of Older Persons.* TRB, National Research Council, Washington, D.C., 1988.
2. L. Evans. *Traffic Safety and the Driver.* Van Nostrand, New York, 1991.
3. M. Edwards. Trends in Women's Fatal Crash Involvement. Presented at the 71st Annual Meeting of the Transportation Research Board, Washington, D.C., 1992.
4. F. Haight. Induced Exposure. *Accident Analysis and Prevention*, Vol. 5, 1973, pp. 111–126.
5. T. Maleck and H. Hummer. Driver Age and Highway Safety. In *Transportation Research Record 1059*, TRB, National Research Council, Washington, D.C., 1987, pp. 6–12.
6. F. McKelvey, T. Maleck, N. Stamatiades, and D. Hardy. Highway Accidents and the Older Driver. In *Transportation Research Record 1172*, TRB, National Research Council, Washington, D.C., 1988, pp. 47–57.
7. N. Garber and R. Srinivasan. Risk Assessment of Elderly Drivers at Intersections. In *Transportation Research Record 1325*, TRB, National Research Council, Washington, D.C., 1991.
8. P. Cooper. Differences in Accident Characteristics Among Elderly Drivers and Between Elderly and Middle-Aged Drivers. *Accident Analysis and Prevention*, Vol. 22, 1990, pp. 499–508.
9. E. Cerelli. Driver Exposure: The Indirect Approach for Obtaining Relative Measures. *Accident Analysis and Prevention*, Vol. 5, 1973, pp. 147–156.
10. R. Lyles, N. Stamatiades, and D. Lighthizer. Quasi-Induced Exposure Revisited. *Accident Analysis and Prevention*, Vol. 23, 1991, pp. 275–285.
11. A. Agresti. *Categorical Data Analysis.* Wiley and Sons, New York, 1990.
12. G. Davis and K. Koutsoukos. Statistical Method for Identifying Locations of High Crash Risk to Older Drivers. In *Transportation Research Record 1375*, TRB, National Research Council, Washington, D.C., 1992.
13. C. Morris. Parametric Empirical Bayes Inference: Theory and Application. *Journal of the American Statistical Association*, Vol. 78, 1983, pp. 47–65.
14. O. Pendleton. *Application of New Accident Analysis Methodologies.* Report FHWA-RD-90-091. FHWA, U.S. Department of Transportation, 1991.
15. J. Maritz. Empirical Bayes Estimation of the Log Odds Ratio in 2 × 2 Contingency Tables. *Commun. Statistics-Theory Meth.*, Vol. 18, 1989, pp. 3215–3233.
16. M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions* (9th ed.). Dover, New York, 1970.

All opinions and conclusions expressed here are solely the responsibility of the authors.

Publication of this paper sponsored by Task Force on Statistical Methods in Transportation.

Comparison of Accident Rates Using the Likelihood Ratio Testing Technique

ALI AL-GHAMDI

Comparing transportation facilities (i.e., intersections and road sections) in terms of traffic accident occurrences is among the interests of most traffic safety analysts. Traditionally, traffic accidents are represented as occurrences of events per certain unit, such as time and vehicle miles; this representation is consistent with Poisson nature. The Poisson distribution is used to describe the distribution of traffic accident occurrences. The objective is to develop a test statistic to enable traffic analysts to compare traffic accident rates in various transportation facilities. The obtained test statistic is simple and requires minimal data to perform the comparison.

Traffic safety improvements have been the concern of traffic engineers lately. The growth in the number of both motorists and traffic accidents is behind this concern. As a result, accident data have been used to analyze traffic accidents as well as to find appropriate techniques to understand the nature of such accidents. This understanding may help traffic accident analysts to draw realistic conclusions regarding the causes or frequencies of accidents and to make appropriate decisions to prevent these causes or reduce the frequencies. This paper uses a hypothesis-testing technique, the likelihood ratio testing technique, to develop a closed form of test statistic to assist traffic analysts in comparing the significance of traffic accident occurrences among different transportation facilities in a transportation system. The occurrences of traffic accidents follow Poisson phenomenon (*1–3*). Hence, the Poisson distribution function is used herein as a basis for deriving the test statistic. The test statistic requires minimal data and can be easily computerized.

STATISTICAL BACKGROUND

Since the approach of this study is based on a statistics technique, the likelihood ratio testing technique, a brief theoretical background of this technique will be given. A general review of statistical distributions and hypothesis testing is given first.

Distributions and Hypothesis Testing

Statistical distributions are useful in interpreting a wide variety of phenomena where randomness is present. In traffic studies the most important distributions are discrete distributions—usually known as counting distributions. Such distributions are useful in describing the occurrence of events that can be counted, such as the number of accidents and the number of arrivals at a certain location.

Two types of discrete distributions are widely used in traffic. They are the binomial and Poisson distributions (*1,3,4*). These distributions have been used in several traffic studies, including studies of speeds, gap acceptances in traffic flow, and accidents. For example, Gerlough and Huber state:

> Counting the number of cars arriving during an interval of time is the easiest and oldest measurement of traffic. When counts from a series of equal time intervals are compared, they appear to form a random series. This led early traffic engineers to investigate distributions as a means of describing the occurrence of vehicle arrivals during an interval. (*1*)

Binomial Distribution

The binomial distribution is formed from a sequence of independent Bernoulli trials, in which the number of successes of a certain number of trials is the quantity of interest. The expansion of the binomial $(q + p)^N$ forms the basis of the binomial distribution function. If N is a positive integer, the $(k + 1)$th term in this expansion is

$$\binom{N}{k} p^k q^{N-k} = \frac{N!}{k(N-k)!} p^k q^{N-k}$$

With parameters N and p, the binomial distribution of a random variable X is defined as

$$\Pr[X = k] = \begin{cases} \binom{N}{k} p^k q^{N-k} & k = 0, 1, 2, \ldots, N \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$ and $q = 1 - p$. The mean and variance of X are Np and Npq, respectively.

This distribution has been used in several traffic applications. In congested traffic (in the case where the ratio of the observed variance/mean is substantially less than 1), for instance, the binomial distribution can be used to describe the distribution of traffic arrivals. When n is very large and p is very small the binomial distribution is approximated by the Poisson distribution.

Poisson Distribution

A random variable X is said to have a Poisson distribution with parameter θ if it has discrete pdf of the form

$$\Pr[X = k] = \frac{e^{-\theta}\theta^k}{k!} \qquad k = 0, 1, 2, \ldots; \theta > 0$$

Department of Civil Engineering, King Saud University, Riyadh, Saudi Arabia.

The random variable X has the same mean and variance. Along with exponential distribution, the pdf of which is defined below, the Poisson distribution has been applied in traffic studies, particularly in studies involving simulation applications. A continuous random variable X has the exponential distribution with parameter $\theta > 0$ if its pdf has the following form:

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

The application of the Poisson distribution to traffic studies has been in existence since the 1930s (*3*). This distribution has been used in fitting traffic accidents and vehicle arrivals at certain locations.

Hypothesis Testing

Hypothesis testing can be defined as the process of making a decision about the truth or the falsehood of a particular hypothesis on the basis of experimental evidence. Generally, experimental outcomes are subject to random error, so any decision made is subject to error too. Occasional decision errors cannot be avoided; however, it is possible to construct tests so that such errors occur infrequently and at some prespecified rate.

For example, suppose our past experience with the traffic accident rate at a specific location indicates that the mean of this rate is 4 if a certain type of traffic control is present, and the mean of such rate may be greater than 4 if that type of control is not present. On the basis of a random sample of size n vehicle accidents in our experiment, we would try to decide which case is true. That is, our test would be the null hypothesis $\mu = 4$ versus the alternative hypothesis $\mu > 4$.

To test a specific hypothesis, a certain critical region is required. The critical region is the subset of the sample space that corresponds to rejecting the null hypothesis. In our example, the sufficient statistic for μ is $\overline{X}$; therefore, we can represent our critical region in terms of the univariate variable—the test statistic. According to the alternative hypothesis, we write our critical region in the following form:

$$C = \{(X_1, \ldots, X_n) | \overline{X} \geq c\}$$

for some appropriate constant c (this constant can be obtained on the basis of the distribution of the random variable in the left-hand side of the inequality). In other words, we will reject the null hypothesis if $\overline{X} \geq c$, and we will accept it if $\overline{X} < c$.

Two possible errors can be made under this testing procedure. The first one is called the Type I Error—rejecting a true H_o. Failing to reject H_o when H_o is false is known as the Type II Error. The objective is to keep both types of error as small as possible. That is, we hope that the selected test statistic and its critical region will yield a small probability of making these two errors. The common notations for these error probabilities are as follows:

$$P[\text{Type I Error}] = \alpha$$

$$P[\text{Type II Error}] = \beta$$

An increase in the sample size will reduce α and β simultaneously. In practice, by selecting a small α we ensure that β will be small too, especially when the sample size is large enough and thus there is no need to specify a value for β. The traditional levels of significance are .01, .05, and .1.

Generalized Likelihood Ratio Test

Suppose $X_1, \ldots, X_n$ have joint pdf $f(x, \theta)$ for $\theta \in \Omega$, and we test the hypothesis $H_o: \theta \in \Omega_o$ versus $H_a: \theta \in \Omega - \Omega_o$. The generalized likelihood ratio (GLR) is defined by

$$\lambda(x) = \frac{\max_{\theta \in \Omega_o} f(x; \theta)}{\max_{\theta \in \Omega} f(x; \theta)} = \frac{f(x; \hat{\theta}_o)}{f(x; \hat{\theta})}$$

where $\hat{\theta}$ is the maximum likelihood estimate (MLE) of θ and $\hat{\theta}_o$ is the MLE under a true H_o. That is, $\hat{\theta}$ and $\hat{\theta}_o$ are determined by maximizing $f(x; \theta)$ over the general parameter space Ω and the restricted parameter space Ω_o. The numerator represents the likelihood function under the null hypothesis (i.e., a subspace of the general parameter space), and the denominator represents the same function but over the general parameter space. The generalized likelihood ratio test is to reject H_o if $\lambda(x) \leq k$, where k is based on the size of significance. In other words, the value of k can be determined to satisfy

$$P[\lambda(x) \leq k | \text{under } H_o] = \alpha$$

It is obvious that if $\lambda(x)$ is a valid statistic (i.e., free of parameters), it will be possible to obtain the exact critical value k. Yet, in many cases the distribution of $\lambda(x)$ is a function of unknown parameters, and thus the critical region cannot be defined. To solve this dilemma, an approximation can take place. That is, MLEs are asymptotically normally distributed. (A distribution dependent on a parameter n, usually a sample number, is said to be asympotatically normal if, as n tends to infinity, the distribution tends to the normal form.) Then it can be proven that the asymptotic distribution of $\lambda(x)$ is free of parameters, and an approximate test will be available to determine the critical region (*5*). In particular, if $x \sim f(X; \theta_1, \ldots, \theta_k)$, then under H_o: $(\theta_1, \ldots, \theta_r) = (\theta_{1o}, \ldots, \theta_{ro})$, $r < k$, for large n, the following approximation holds:

$$-2 \log \lambda(x) \sim \chi^2_r$$

Thus, H_o is rejected if

$$-2 \log \lambda(x) \geq \chi^2_{1-\alpha, r}$$

ANALYSIS

The analysis in this study consists of two stages. First, real accident data were used for four types of highways in Ohio (Table 1) to develop a test statistic for comparing their accident rates. This test statistic was generalized to be applicable for different types of data.

TABLE 1 Accident Rate Data for Ohio (*6*)

Highway Type	*Characteristics*		Accident Rate
	All Accidents	AMVM	
Scenic	3,621	1,021	3.55
Other 2-lane	36,752	11,452	3.21
Multi-lane	20,348	6,290	3.23
Interstate	10,460	9,412	1.11
Total	71,181	28,177	2.53

Derivation of the Test Statistic Based on Real Data

Accident data from the Ohio Department of Transportation are shown in Table 1 (*6*). The table presents 1-year accidents for different highway types, including scenic, other 2-lane, multilane, and Interstate. The last column of the table presents the accident rate for each type. This rate is the total number of accidents divided by the annual million vehicle miles (AMVM).

Our interest is to find out whether accident rates among these types are different. In other words, the numbers suggest some differences among accident rates, but the question can be asked whether such differences are true differences or the result of randomness.

Since we are dealing with the occurrence of number of accidents (events) per AMVM (unit of exposure), it is worthwhile to assume that the number of accidents (X) is Poisson distributed. Thus X_i is Pois(μ_i), $i = 1, 2, 3, 4$ (i represents a highway type), and

$$f(x_i, \mu_i) = \frac{e^{-\mu_i}\mu_i^{x_i}}{x_i!}$$

where $x_i = 0, 1, 2, \ldots$; $i = 1, 2, 3, 4$; and $\mu_i > 0$. In addition,

$$\mu_i = \lambda_i t_i$$

where λ_i is the accident rate for highway type i and t_i is the annual vehicle miles for highway type i.

To test whether such rates are unequal, we need to develop our hypotheses. The null and alternative hypotheses are H_o: $\lambda_1 = \lambda_2 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda$ and H_a:$\lambda_i \neq \lambda_k$ for some i, k (i and k are two different types of highways).

The likelihood ratio technique is used to test the above hypothesis. The joint pdf of X_1, X_2, X_3, and X_4, also called the likelihood function, is

$$L = \prod_{i=1}^{4} \frac{(\lambda_i t_i)^{x_i}}{x_i!} e^{-\lambda_i t_i} \tag{1}$$

By taking the log of Equation 1, it can be simplified to

$$\log L = \sum_{i=1}^{4} [x_i \log(\lambda_i t_i) - \log(x_i!) - \lambda_i t_i]$$

$$= \sum_{i=1}^{4} [x_i \log \lambda_i + x_i \log t_i - \log(x_i!) - \lambda_i t_i] \tag{2}$$

Under H_o the derivative of Equation 2 is obtained, set to equal 0, and solved for the parameter λ:

$$\frac{\partial \log L}{\partial \lambda} = \frac{\sum_{i=1}^{4} x_i}{\lambda} - \sum_{i=1}^{4} t_i \overset{\text{set}}{=} 0$$

$$\rightarrow \frac{\sum_{i=1}^{4} x_i}{\lambda} = \sum_{i=1}^{4} t_i$$

$$\rightarrow \hat{\lambda} = \frac{\sum_{i=1}^{4} x_i}{\sum_{i=1}^{4} t_i} = \frac{\bar{x}}{\bar{t}} \tag{3}$$

where

$$\bar{x} = \frac{\sum_{i=1}^{4} x_i}{4}$$

and

$$\bar{t} = \frac{\sum_{i=1}^{4} t_i}{4}$$

This is the MLE (this solution maximizes the likelihood function under the null hypothesis). Thus, under H_o the maximum of likelihood function equals

$$L_o = \prod_{i=1}^{4} \frac{(\hat{\lambda} t_i)^{x_i}}{x_i!} e^{-\hat{\lambda} t_i}$$

$$= \prod_{i=1}^{4} \frac{\left(\frac{\bar{x} t_i}{\bar{t}}\right)^{x_i}}{x_i!} e^{-(\bar{x} t_i/\bar{t})}$$

$$= \exp\left(-\frac{\bar{x} \sum_{i=1}^{4} t_i}{\bar{t}}\right) \prod_{i=1}^{4} \frac{\left(\frac{\bar{x}_i}{\bar{t}}\right)^{x_i}}{x_i!}$$

$$= e^{-4\bar{x}} \prod_{i=1}^{4} \frac{\left(\frac{\bar{x} t_i}{\bar{t}}\right)^{x_i}}{x_i!} \tag{4}$$

Equation 4 will be used, shortly, as the numerator in the likelihood ratio function. Over the general parameter space where H_a:$\lambda_i \neq \lambda_k$ for some i and k, by taking the derivative of Equation 2 with respect to each λ_i we obtain

$$\frac{\partial \log L}{\partial \lambda_1} \overset{\text{set}}{=} 0 \Rightarrow \hat{\lambda}_1 = \frac{x_1}{t_1}$$

$$\frac{\partial \log L}{\partial \lambda_2} \overset{\text{set}}{=} 0 \Rightarrow \hat{\lambda}_2 = \frac{x_2}{t_2}$$

$$\frac{\partial \log L}{\partial \lambda_3} \overset{\text{set}}{=} 0 \Rightarrow \hat{\lambda}_3 = \frac{x_3}{t_3}$$

$$\frac{\partial \log L}{\partial \lambda_4} \overset{\text{set}}{=} 0 \Rightarrow \hat{\lambda}_4 = \frac{x_4}{t_4} \tag{5}$$

which is the MLE under the alternative hypothesis. Under the alternative hypothesis, where λ_i's are not the same, the maximum of likelihood function becomes

$$L_a = \prod_{i=1}^{4} \frac{\left(\frac{x_i}{t_i} t_i\right)^{x_i}}{x_i!} e^{-(x_i/t_i)} t_i$$

$$= e^{-4\bar{x}} \prod_{i=1}^{4} \frac{(x_i)^{x_i}}{x_i!} \tag{6}$$

The ratio of L_o to L_a is called the likelihood ratio and is denoted by

$$\psi = \frac{L_o}{L_a} = \frac{\prod_{i=1}^{4} \frac{(\hat{\lambda} t_i)^{x_i}}{x_i!} e^{-\hat{\lambda} t_i}}{e^{-4\bar{x}} \prod_{i=1}^{4} \frac{(x_i)^{x_i}}{x_i!}}$$

$$= \frac{\prod_{i=1}^{4} \left(\frac{\bar{x} t_i}{\bar{t}}\right)^{x_i}}{\prod_{i=1}^{4} (x_i)^{x_i}}$$

$$= \prod_{i=1}^{4} \left(\frac{\frac{\bar{x} t_i}{\bar{t}}}{x_i}\right)^{x_i}$$

$$= \prod_{i=1}^{4} \left(\frac{\bar{x} t_i}{x_i \bar{t}}\right)^{x_i} \tag{7}$$

Hence, the test statistic, approximately, for large n is $-2 \ln \psi$, which has chi-square distribution. Specifically,

$$-2 \ln \psi = -2 \sum_{i=1}^{4} x_i[\ln(\bar{x} t_i) - \ln(\bar{t} x_i)]$$

$$= -2\left[\sum_{i=1}^{4} x_i \ln(\bar{x} t_i) - \sum_{i=1}^{4} x_i \ln(\bar{t} x_i)\right] \tag{8}$$

is chi-square with three degrees of freedom, and the approximate size test is to reject H_o if

$$-2 \ln \psi \geq \chi^2_{1-\alpha}(4 - 1)$$

Thus, the critical region for the above test can be defined through the following form:

$$P[-2 \ln \psi \geq \chi^2_{1-\alpha}(4 - 1)] = \alpha$$

Generalization of the Test Statistic

The test statistic reached in the solution of the data given in the previous section can be generalized to cover more applications as long as the setup for Table 1 is unchanged. The general setup for this table is presented in Table 2.

In this table the unit of exposure could be any type of units used when accidents were observed, such as vehicle miles, hours, or days. The facility type refers to the place where the accidents take place. In the field of transportation people are served by a variety of facilities, including highways, intersections, local streets, and parking lots. The variables listed in this table are as follows:

TABLE 2 Accident Rate Data, General Setup

Highway Facility Type	*Characteristics*		Accident Rate
	All Accidents	Unit of Exposure	
1	x_1	t_1	λ_1
2	x_2	t_2	λ_2
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
j	x_j	t_j	λ_j
Total	$\sum_{i=1}^{j} x_i$	$\sum_{i=1}^{j} t_i$	

x_i = the total number of accidents that occur in Facility i,
t_i = the unit during which accidents occur in Facility i (exposure), and
λ_i = the accident rate at Facility $i = x_i/t_i$.

Notice that the analyst could use any unit of exposure on the basis of the available data.

The variable of interest x_i is assumed to have Poisson distribution with parameter λ_i. Thus, the test statistic derived in the previous section can be slightly modified to take the following general form:

$$-2 \ln \psi = -2\left[\sum_{i=1}^{j} x_i \ln(\bar{x} t_i) - \sum_{i=1}^{j} x_i \ln(\bar{t} x_i)\right] \tag{9}$$

where

$\bar{x}$ = mean of x_i's, $i = 1, 2, \ldots, j$;
$\bar{t}$ = mean of t_i's;
i = Facility i; and
j = number of facilities under consideration.

This test statistic can be used to detect the difference among accident rates for any type of facilities provided that the above table setup is satisfied. This test statistic is chi-square distributed with $j - 1$ degrees of freedom. Notice that this test statistic requires only the number of accidents occurring at Facility i and the desired unit of exposure during which these accidents occur. This general form is applicable for any number of transportation facilities (j).

Recall $\lambda_j = x_j/t_j$. Accident rates, in literature, are usually compared in terms of their quantities. In other words, of two locations, the one with the higher accident rate is considered to be more severe. Unfortunately, this is misleading. That is, it may be inconclusive statistically since the difference may be due to chance. The test statistic developed in this paper, however, can detect whether such a difference is significant or due to chance.

APPLICATION

In the previous sections we went step by step through the likelihood ratio technique to test the hypothesis of equal ac-

TABLE 3 Pairwise Comparisons for the Four Highway Types Presented in Table 1

The Highway Pair*	Significance at 5% level
1&2	significant
1&3	significant
1&4	significant
2&3	insignificant
2&4	significant
3&4	significant

* The numerical codes: 1 for Scenic, 2 for Other 2-lane, 3 for Multi-lane, 4 for Interstate.

cident rates given in Table 1, and we ended with the general form of a test statistic defined in Equation 9 to perform the hypothesis testing. In this section we apply this test statistic to the data given in that table. Moreover, a computer program was written to perform the computations of the test statistic.

In Table 1, four types of highways are presented. Therefore, j is 4 in our general form and we have the following hypothesis test: H_o: $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda$ and H_a: $\lambda_i \neq \lambda_k$ for some i, k. The test statistic in Equation 9 is

$$-2 \ln \psi = -2\left[\sum_{i=1}^{4} x_i \ln(\bar{x}t_i) - \sum_{i=1}^{4} x_i \ln(\bar{t}x_i)\right]$$

A comparison of the four groups indicated that the value of the test statistic is greater than 7.81 (the critical value at 0.05 level of significance), indicating that the null hypothesis is rejected and the difference among the accident rates for these four highway types is significant. In fact, this result was expected, since Table 1 indicates such differences, particularly between scenic highway and Interstate highway types.

If we decide to make pairwise comparisons, different results are obtained. For example, when we compare other two-lane with multilane ($j = 2$ in this case), the value of the test statistic is very small, namely 0.9375, which in turn means that the difference between accident rates for these two types is not significant. Table 3 presents the pairwise comparisons of the highway types given in Table 1. Figure 1 shows the pairwise comparisons. The insignificant pair is underlined in Figure 1.

Type 1	Type 2	Type 3	Type 4
1	2	3	4

FIGURE 1 Graphical representation of pairwise comparisons presented in Table 3.

CONCLUSION

The finding of this paper was a test statistic for the comparison of accident rates in several transportation facilities. This finding was based on the assumption that such accidents were Poisson distributed. The likelihood ratio statistical technique was used to develop the test statistic. With minimal data this statistic can be adopted by traffic analysts to detect whether accident rates at several locations in a transportation system are significantly different. To show the applicability of the derived test statistic, traffic accident data from Ohio were used to compare accident rates for four highway types, including scenic, other two-lane, multilane, and Interstate. These rates were found to be significantly different. Pairwise comparisons for these types indicated that there is no significant difference between the accident rates for the other two-lane type and the multilane type. The results of this study have shown the applicability of the developed test statistic.

ACKNOWLEDGMENTS

The author wishes to thank the officials of the Ohio Department of Transportation, Columbus, Ohio, for their help in making available the data used in this study. The valuable suggestions and comments made by Ramey Rogness of the Department of Civil Engineering at The Ohio State University during the course of this research are also appreciated.

REFERENCES

1. Gerlough and Huber. *Statistics with Applications to Highway Traffic Analyses.* Eno Foundation for Transportation, Inc., Westport, Conn., 1978.
2. Gerlough and Huber. *Traffic Flow Theory: A Monograph.* TRB, National Research Council, Washington, D.C., 1975.
3. Gerlough and Barnes. *Poisson and Other Probability Distributions in Highway Traffic.* Eno Foundation for Transportation, Inc., Westport, Conn., 1971.
4. Taylor and Young. *Traffic Analysis: New Technology New Solutions.* Hargreen Publishing Co., Maryborough, Victoria, 1988.
5. Lee and Engelhardt. *Introduction to Probability and Mathematical Statistics.* Duxbury Press, Boston, 1987.
6. *Highway Safety Improvement Programs: Progress and Evaluation Report, Fiscal Year 1990.* Ohio Department of Transportation, Sept. 1990.

Publication of this paper sponsored by Task Force on Statistical Methods in Transportation.

Accident Prediction Models for Freeways

BHAGWANT PERSAUD AND LESZEK DZBIK

The modeling of freeway accidents continues to be of interest because of the frequency and severity of these accidents and the congestion associated with them. Some difficulties with conventional modeling techniques are identified. A distinctive approach is presented, whereby generalized linear modeling is used with both macroscopic and microscopic data to develop regression model estimates of a freeway section's accident potential and an empirical Bayesian procedure is used for refining these estimates.

Freeway accidents are a source of concern not only because of their frequency and severity but also because of the resulting traffic congestion. It is, therefore, not surprising that attention continues to be focused on the modeling of these accidents to identify associated factors and to enable analysts to predict their frequency. Recent papers (*1–3*) are evidence of this ongoing interest. This paper is based on recent research (*4,5*) that applied an accident modeling approach that is somewhat distinct from those used by others.

FREEWAY ACCIDENT MODELING ISSUES

In this section, a number of difficulties with previous models are reviewed. The approach adopted for this paper is then introduced.

The first difficulty with existing models is that they tend to be macroscopic in nature since they relate accident occurrence to average daily traffic (ADT) rather than to the specific flow at the time of accidents. The difficulty with the macroscopic approach is that a freeway with intense flow during rush periods would clearly have a different accident potential than a freeway with the same ADT but with flow evenly spread out during the day, but an ADT-based model would indicate that the two freeways have identical accident potentials.

Second, some modelers assume, a priori, that accidents are proportional to traffic volume and go on to use accident rate (accidents per unit of traffic) as the dependent variable. There is much research to suggest that this assumption is not only incorrect but can also lead to paradoxical conclusions (*6*). Similarly, though accidents should increase with traffic intensity, the model form should not, a priori, assume that accidents are a linear function of traffic volume (*7*).

Third, conventional regression modeling assumes that the dependent variable has a normal error structure. For accident counts, which are discrete and nonnegative, this is clearly not the case; in fact, a negative binomial error structure has been shown to be more appropriate (*8*). Most regression packages in use cannot accommodate such a structure.

Finally, it is impossible for regression models to account for all of the factors that affect accident occurrence. This difficulty can lead to paradoxical conclusions when, as is often done, such models are used to imply cause and effect. Also, when these models are used for accident prediction, the estimates tend to be unreliable if the unexplained variation is relatively large.

The need to overcome these difficulties was fundamental to the modeling approach adopted in the work described in this paper. To this end, use was made of a generalized linear modeling package that allows the flexibility of a nonlinear accident-traffic relationship and a user-specified error structure for the dependent variable and of a complementary empirical Bayesian procedure for improving the accuracy of regression model accident predictions. The approach was applied to both microscopic data (hourly accidents and hourly traffic) and macroscopic data (yearly accident data and average daily traffic).

THEORETICAL ASPECTS OF REGRESSION MODELING

Generalized linear modeling using the GLIM computer package (*9*) was used to obtain a regression model for estimating P, the accident potential per kilometer per unit of time, given a freeway section's physical characteristics, the volume (T) per unit of time, and a set of variables that describe operating conditions during the time period. The model form used was

$$E(P) = aT^b \tag{1}$$

where a and b are model parameters estimated by GLIM. Models were so constructed that the parameters a and b could depend on the values of the factorial variables. This model form ensures that predicted accidents would be zero if there is no traffic, but does not, a priori, assume a linear relationship between accidents and traffic volume. Scatter plots of raw data confirmed that this model form is reasonable for both macroscopic and microscopic models.

The accident count on a section was used as an estimate of the dependent variable. GLIM allows the specification of a negative binomial error structure for a dependent variable, which, as noted earlier, is more appropriate for accident counts than the traditional normal distribution. Although the error structure pertains to the accident counts, a log link function could be specified to allow GLIM to estimate models of the form

$$\ln[E(P)] = \ln(a) + b\ln(T) \tag{2}$$

B. Persaud, Department of Civil Engineering, Ryerson Polytechnical Institute, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada. L. Dzbik, Department of Civil Engineering and Engineering Mechanics, McMaster University, Hamilton, Ontario L8S 4L7, Canada.

With a negative binomial error specification, it can be shown that the variance of the regression estimates can be estimated from

$$\mathrm{Var}(P) = \mathrm{E}(P)^2/k \quad (3)$$

where the parameter k was estimated using a maximum likelihood procedure that assumes that each squared residual of the regression model is an estimate of Var(P) and that each count comes from a negative binomial distribution with mean E(P) and variance given by Equation 3. This equation indicates that, in comparing two models with the same dependent variable, the one with the larger value of k would give more accurate predictions.

Because ordinary least squares regression was not used, goodness-of-fit of a model could not be assessed in the conventional way, using the coefficient of determination. Instead, goodness-of-fit was assessed by using a generalized Pearson chi-squared statistic (*8,9*) to estimate the amount of variation explained by the systematic component of a model.

MACROSCOPIC MODELS

Data

The data originated in computer files obtained from the Ontario Ministry of Transportation and consists of accident, inventory, and traffic data for approximately 500 freeway sections in Ontario. Some characteristics of the macroscopic data set are summarized in Table 1.

Model Calibration and Results

For each section, the accident count for the each of the years 1988 and 1989 (in effect, the log of this value) was used as an estimate of the dependent variable. To account for varying section lengths, the term log (section length) was specified as an "offset" that GLIM subtracts from each point estimate of ln[E(P)]. Thus, in effect, models were estimated for prediction of the number of accidents per kilometer per year.

Tables 2 and 3 give the estimated regression model coefficients for total accidents and severe (injury and fatal) accidents.

MICROSCOPIC MODELS

Data

The data pertain to a 25-km segment of Highway 401 in Toronto, Canada, part of which has a Freeway Traffic Management System (FTMS). The sections, which range in length from 0.7 to 3 km, are separated by interchanges, and all have express and collector roadways typically with three lanes each per direction.

For the microscopic modeling, it was necessary that conditions pertaining to each data record used in the regression analysis be fairly homogeneous. Thus, it was decided to disaggregate each day into 24 periods of 1 hr each and to derive data for each hour, for express and collector lanes separately, and for day and night. For the accident data this task was straightforward. For the traffic data, it was necessary to derive hourly and seasonal variation factors and collector/express lane distribution factors and apply these factors to the average daily traffic. To maintain a reasonable level of homogeneity, only data pertaining to weekdays were used for the models presented in this paper.

After preliminary data analysis that indicated different accident patterns for congested and uncongested periods, it was decided to build the regression models using, for each section, only data for off-peak hours for which that section tended to be uncongested. To make this determination, we used 5 days of traffic data for sections in the FTMS and used a procedure described elsewhere (*10*). Congested and uncongested hours for a section were identified as hours for which the applicable condition existed on all 5 days. It was assumed that any errors in this process would have a negligible effect since the amount of incorrectly classified data was likely to be relatively small.

TABLE 1 Data Summary for Macroscopic Models

	4-lane	> 4 lanes
Total km	1594	397
1988-89 total accidents	13725	24464
1988-89 severe accidents	4999	7519
ADT (Weighted by length)	19621	87896

TABLE 2 Macroscopic Model for TOTAL Accidents per Kilometer per Year

Model Parameter	Estimated Parameter Adjustment	Standard Error*	Adjusted Parameter Estimate
ln(a) for ADT/1000:			
4 lanes	0 (Base)	0.087	-1.920
> 4 lanes	0.271	0.062	-1.649
b for:			
all lanes	0 (Base)	0.028	1.135
k = 3.52; Variation Explained = 98%; Observations = 1012			

* Applies to coefficient estimate for the base case; otherwise, applies to the adjustments.

TABLE 3 Macroscopic Model for SEVERE Accidents per Kilometer per Year

Model Parameter	Estimated Parameter Adjustment	Standard Error	Adjusted Parameter Estimate
ln(a) for ADT/1000:			
4 lanes	0 (Base)	0.126	-2.776
> 4 lanes	-0.417	0.254	-3.193
b for:			
4 lanes	0 (Base)	0.040	1.082
> 4 lanes	0.124	0.068	1.206
k = 4.55; Variation Explained = 93%; Observations = 1012			

The regression data set contained, for each section and each uncongested hour, the hourly average traffic volume, the number of applicable hours in the 2-year period 1988–1989, a tally for each accident type of interest for those hours, and a code to indicate the light condition (day/night). For some hours (e.g., 6:00 to 7:00 p.m.), it was necessary to have separate sets of data for day and night conditions. A summary of information in the regression data set is given in Table 4.

Model Calibration and Results

For each section, the accident count for the 2-year period 1988–1989 for each uncongested hour (in effect, the log of this value) was used as an estimate of the dependent variable. To account for varying section lengths and number of hours of data, the term ln(section length * number of hours) was specified as an "offset" that GLIM subtracts from each point estimate of ln[E(P)]. Thus, in effect, models were estimated for prediction of the number of accidents per kilometer per hour.

Tables 5 and 6 give the estimated regression model coefficients for total accidents and severe (injury and fatal) accidents. In the calibration process it was found that there was no significant difference between day and night accident frequencies.

As indicated, the estimated coefficients are for a 1-km section for 1 hr. Thus, the regression estimate of total accident potential for a 2-km collector section during an hour with a volume of, say, 8,000 vehicles, is given by $E(P) = 2*e^{-6.276}*8^{0.717} = 0.01671$ accidents/hour and $Var(P) = 0.01671^2/2.59 = 0.000108$.

DISCUSSION OF REGRESSION MODEL RESULTS

Figure 1 shows plots of regression predictions per kilometer per year for the macroscopic models. These plots indicate

TABLE 4 Data Summary for Microscopic Models

	Hourly Volume		Day	Night	Collector Accidents				Express Accidents			
					Severe		Total		Severe		Total	
Hr	Collector	Express	Hrs	Hrs	Day	Night	Day	Night	Day	Night	Day	Night
00	1542	2167	-	521	-	17	-	46	-	14	-	31
01	802	1239	-	521	-	6	-	29	-	9	-	20
02	465	848	-	521	-	6	-	23	-	5	-	19
03	351	646	-	521	-	1	-	11	-	0	-	11
04	446	815	-	521	-	2	-	6	-	3	-	15
05	1382	2128	-	325	-	1	-	3	-	4	-	11
06	6642	9196	44	127	2	2	9	21	2	5	6	16
07	10477	12825	264	-	32	-	101	-	24	-	78	-
08	10316	12015	521	-	73	-	218	-	46	-	177	-
09	7946	9860	521	-	34	-	127	-	31	-	108	-
10	6708	9042	521	-	32	-	111	-	24	-	83	-
11	7012	9023	521	-	33	-	108	-	30	-	84	-
12	6944	8858	521	-	27	-	84	-	20	-	71	-
13	7357	9330	521	-	28	-	89	-	52	-	110	-
14	8153	10062	521	-	45	-	119	-	27	-	106	-
15	9826	11543	521	-	54	-	185	-	60	-	180	-
16	10538	11950	391	-	78	-	231	-	68	-	212	-
17	10047	11042	307	43	65	6	192	28	58	8	161	29
18	8686	10334	220	171	25	24	84	81	23	16	64	65
19	6425	8103	198	260	13	21	34	66	9	20	20	75
20	4683	6126	-	151	-	12	-	32	-	14	-	56
21	4463	5545	-	521	-	26	-	68	-	21	-	74
22	3793	4821	-	521	-	21	-	74	-	27	-	104
23	3057	3845	-	521	-	25	-	55	-	35	-	89
TOTALS					541	170	1692	543	474	181	1460	615

TABLE 5 Microscopic (Off-Peak) Model for TOTAL Accidents per Hour per Kilometer

Model Parameter	Estimated Parameter Adjustment	Standard Error*	Adjusted Parameter Estimate
ln(*a*) for Vol/Hr/1000:			
Collector	0 (Base)	0.081	-6.276
Express	-0.258	0.077	-6.534
***b* for:**			
Collector/Express	0 (Base)	0.045	0.717
k = 2.59; Variation Explained = 87%; Observations = 684			

* Applies to coefficient estimate for the base case; otherwise, applies to the adjustments.

TABLE 6 Microscopic (Off-Peak) Model for SEVERE Accidents per Hour per Kilometer

Model Parameter	Estimated Parameter Adjustment	Standard Error	Adjusted Parameter Estimate
ln(*a*) for Vol/Hr/1000:			
Collector	0 (Base)	0.116	-7.608
Express	-0.276	0.098	-7.883
***b* for:**			
Collector/Express	0 (Base)	0.063	0.777
k = 2.30; Variation Explained = 87%; Observations = 684			

that, for the same total traffic volume, four-lane freeways have a lower accident risk than those with more lanes. This result is possibly explained by the tendency for freeways with more than four lanes to be found in urban areas that are generally associated with rush hour congestion and an accompanying greater collision risk.

Figure 2 shows plots of microscopic model regression predictions per kilometer per hour for the two accident types and for express and collector roadways. It is evident that, for a given traffic volume level, collector roadways have a higher accident potential than the express roadways. It is important to note that, for these regression lines, the slope is decreasing as hourly volume increases, perhaps capturing the influence of decreasing speed. This is in contrast to the macroscopic plots in Figure 1, which all show increasing slopes. It is possible that the macroscopic plots are reflecting the increasing probability of risky maneuvers, such as passing and lane changing, with higher ADT levels.

The issue of how accident risk is related to the quality of traffic operation was examined separately. Recall that the microscopic regression models were based on data for hours without congestion. For congested hours, the average hourly traffic volume was calculated along with the average number of accidents per hour per kilometer. Separate values were calculated for the collector and express systems, and in the case of total accidents it was possible to calculate separate

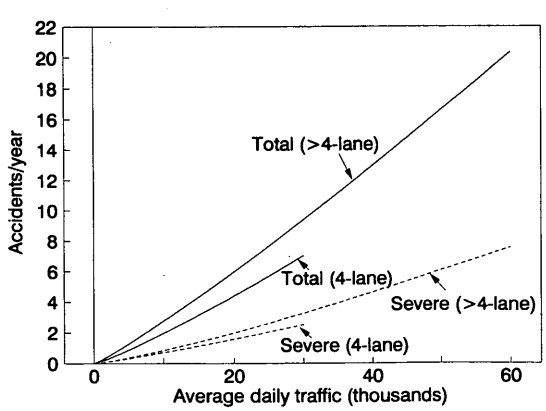


FIGURE 1 Macroscopic regression model predictions.

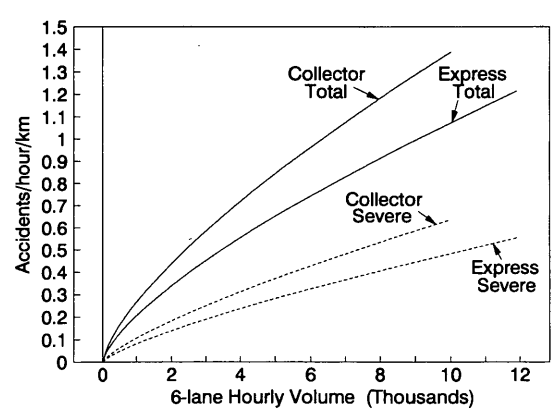


FIGURE 2 Microscopic regression model predictions.

values for morning and evening congested periods. This supplementary work was exploratory in nature, and the results are to be interpreted with caution since they are based on untested assumptions, in particular about the time and location of congestion and about the propriety of using an average traffic volume for congested periods. The results of the analysis of the effect of traffic operation are shown in Figures 3 and 4. Subject to the cautions mentioned, the following conclusions are indicated:

- Congestion is associated with a higher risk of accidents than high-volume uncongested operation.
- The afternoon congested period has a higher accident risk than the morning rush period, but the difference is only significant for the express system.
- Collector system congestion is associated with a higher accident risk than express system congestion.

ACCIDENT PREDICTION—REFINEMENT OF REGRESSION MODEL ESTIMATES

It is now accepted among safety analysts that the underlying long-term accident potential, rather than the commonly used short-term count, is more proper for identifying unsafe sections and for evaluating safety effectiveness of improvements. Regression model predictions have been used as an estimate of this value, but the difficulty with this is that, in general, two road sections that are similar in all of the independent variables used in a regression model will still be different in true accident potential even though they will have the same model predictions. This, in turn, is because it is not possible to account in the regression model for all the factors that cause differences in accident potential (e.g., weather, geometrics).

To mitigate this problem, use can be made of an empirical Bayesian technique that combines the regression prediction with the observed short-term accident count for a section of interest. This method has previously been used in estimating the long-term accident potential of rail-highway grade crossings (*11*), Toronto intersections (*12*), and Ontario drivers (*13*) and road sections (*4*).

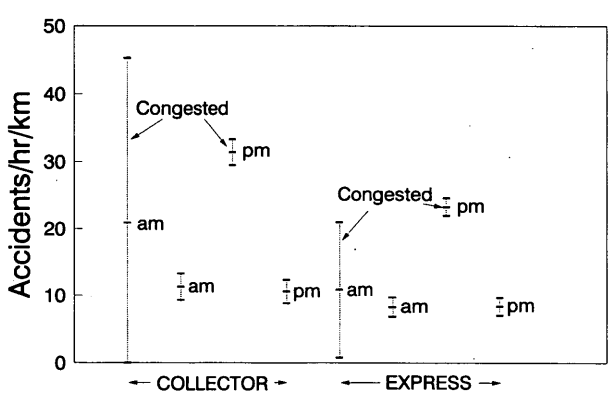


FIGURE 3 Point estimates (with 95 percent confidence intervals) of accident potential during high-volume operation—TOTAL ACCIDENTS.

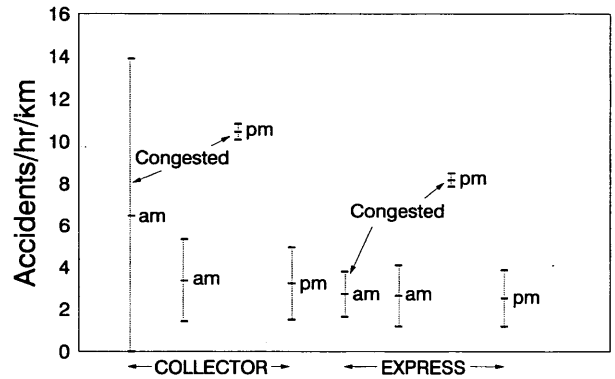


FIGURE 4 Point estimates (with 95 percent confidence intervals) of accident potential during high-volume operation—SEVERE ACCIDENTS.

Theory

Using the empirical Bayesian procedure, E(P) from Equation 1 can be refined for an individual road section using the accident count, x, in n units of time (years in the case of macroscopic models and hours for the microscopic case) on that section to give E($m|P, x, n$), a revised estimate of accident potential. It can be shown (*11*) that, under reasonable assumptions, the revised estimate of accident potential (per unit of time) is

$$E(m|P, x, n) = q[wE(P) + (1 - w)x] \quad (4)$$

where

E(P) = regression estimate for one unit of time,
$w = [1/(1 + \mathrm{Var}(P)/E(P))] = [1/(1 + E(P)/k)]$, and
$q = [(1 + E(P)/k)/(1 + nE(P)/k)]$.

It can also be shown that the variation in ($m|P, x, n$) can be estimated by

$$\mathrm{Var}(m|P, x, n) = \{E(m|P, x, n)/[n + k/E(P)]\} \quad (5)$$

Equation 4 indicates that the estimated accident potential of a section is a combination of what is observed, x, and of E(P)—what is predicted on the basis of its characteristics (traffic volume, etc.).

Illustration

Suppose the collector section in the earlier example recorded six accidents in 80 hr with an average hourly traffic volume of 8,000. Thus, in Equation 4, $x = 6$ and $n = 80$. Recall from the earlier example that E(P) = 0.01671, Var(P) = 0.000108, and $k = 2.59$. Thus, for Equations 4 and 5, $q = 0.6638$ and $w = 0.9936$.

These values give the Bayesian estimate of accident potential as E($m|P$, 6, 80) = 0.03651 accidents per hour and Var($m|P$,

TABLE 7 Macroscopic Model Validation Results—Mean Squared Difference Between Predicted and Observed Accidents per Kilometer per Year

Estimation method	Total accidents	Severe accidents
Accident count	32.2	5.94
Regression model	31.9	5.84
Empirical Bayesian	23.0	3.78

TABLE 8 Microscopic Model Validation Results—Mean Squared Difference Between Predicted and Observed Accidents per Hour-Kilometer

Estimation method	Total accidents ($*10^{-6}$)	Severe accidents ($*10^{-6}$)
Accident count	0.928	3.573
Regression model	0.771	2.682
Empirical Bayesian	0.743	2.469

6, 80) = 0.03651/(80 + 2.59/0.01671) = 0.000155. Note that the accident potential estimate is between the regression estimate (0.01671) and the observed accidents per hour (6/80 = 0.075).

Validation

The validation of the overall approach involved a comparison of the prediction accuracy resulting from the use of 1988 accident counts or regression predictions as estimates of the 1987 counts, as opposed to using the empirical Bayesian procedure based on 1988 data.

For the macroscopic case, this required the calculation of the mean squared difference between estimated and observed 1987 total and severe accidents for each of approximately 1500 km of freeways for each of the three estimation methods. It is assumed that the better estimate is the one with the smallest mean squared difference. The results of the validation exercise for the macroscopic models are given in Table 7.

For the microscopic case, validation required the calculation of the squared difference between estimated and observed 1987 counts per squared hour-kilometer for each cell (see Table 4) for each section, and, in essence, averaging this value over all sections and cells. The results of the validation exercise for the microscopic models are given in Table 8.

The results in both cases show that the empirical Bayesian method appears to be best followed, as expected, by the regression model prediction method.

ACKNOWLEDGMENTS

The research for this paper was supported under an operating grant from the Natural Sciences and Engineering Council of Canada. The assistance of the Ministry of Transportation of Ontario in providing the data and in funding related research is greatly appreciated.

REFERENCES

1. Y. Huang, R. Cayford, and A. D. May. Accident Prediction Models for Freeway Segments. Presented at 71st Annual Meeting of the Transportation Research Board, Washington, D.C., 1992.
2. Kraus et al. Epidemiological Aspects of Fatal and Severe Injury Urban Freeway Crashes. *Accident Analysis and Prevention* (forthcoming).
3. A. Ceder and M. Livneh. Relationships Between Road Accidents and Hourly Traffic Flow—I. Analysis and Interpretation. *Accident Analysis and Prevention*, Vol. 14, No. 1, 1982, pp. 19–34.
4. B. N. Persaud. *Accident Potential Models for Ontario Road Sections*. Ministry of Transportation of Ontario, March 1991.
5. L. Dzbik. *Accident Prediction Models for Freeways*. Master's thesis. McMaster University, Hamilton, Ontario, Canada, 1992.
6. D. Mahalel. A Note on Accident Risk. In *Transportation Research Record 1068*, TRB, National Research Council, Washington, D.C., 1985.
7. S. P. Satterthwaite. *A Survey of Research into the Relationships Between Traffic Accidents and Traffic Volumes*. Transport and Road Research Laboratory Supplementary Report 692, United Kingdom, 1981.
8. E. Hauer, J. Lovell, and B. N. Persaud. *New Directions for Learning About Safety Effectiveness*. Report FHWA/RD-86-015. Federal Highway Administration and National Highway Traffic Safety Administration, Jan. 1986.
9. R. J. Baker and J. A. Nelder. *The GLIM System—Release 3*. Rothamsted Experimental Station, Harpenden, United Kingdom, 1978.
10. L. Aultman-Hall and F. L. Hall. *Demonstration of a Method Using Only Detector Data To Evaluate the Effectiveness of Highway 401 FTMS*. Department of Civil Engineering, McMaster University, Nov. 1991.
11. E. Hauer and B. N. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effect of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1986.
12. E. Hauer, J. C. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988.
13. A. Smiley, B. N. Persaud, E. Hauer, and D. Duncan. Accidents, Convictions and Demerit Points—An Ontario Driver Records Study. In *Transportation Research Record 1238*, TRB, National Research Council, Washington, D.C., 1989, pp. 53–64.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Use of Freeway Conflict Rates as an Alternative to Crash Rates in Weaving Section Safety Analyses

JOSEPH FAZIO, JANET HOLDEN, AND NAGUI M. ROUPHAIL

Traffic safety is an important concept in the evaluation of a transportation system and its impact on public health. Using reported crash rates as an indicator of safety of a freeway facility has many drawbacks, such as errors in the reporting and recording of crashes, inaccuracies in the way in which the exposure measure is derived, and the wait involved for a sufficient sample size to materialize. Conflict rates provide an alternative to crash rates as an indicator of safety. Benefits of their use include the ease and accuracy with which conflict rates at ramp weaves can be obtained and the high frequency at which they occur with no physical harm to the public. Computer subroutines were added to the Integrated Traffic Simulation (INTRAS) to count conflicts, and freeway traffic was simulated at 10 modeled ramp weaves on Interstate 294 (Interstate 294 serves as a quasi-beltway for the Chicago metropolitan area). The resulting conflict rates were then noted. Volume, geometric, and crash data for the 10 sites were provided by the Illinois State Toll Highway Authority and on-site visits. The conflict rates at the sites were applied in a test of their ability to identify the known hazardous ramp weaves, and the relationship between conflict and crash rates was examined.

Weaving is the crossing of traffic streams flowing in the same general direction. Freeway weaving sections have long been a problem area for highway engineers in terms of their design, traffic operations analysis, and safety. When a freeway driver performs a weaving maneuver in the weaving section, the driver interacts with entering and exiting drivers possibly in a conflicting manner. A traffic conflict is an event that has the potential of being a traffic crash.

Studies on freeway traffic conflicts are few in number. The main reason for this discrepancy is the high average crash rates observed on the nonfreeway system. The crash frequencies on freeways are far less than the crash frequencies on roadways with at-grade intersections. For instance, in 1988 Illinois urban Interstate highway crashes totaled 32,546, whereas urban non-Interstate highway crashes totaled 426,046 (*1*,p.11). Urban includes locations in or adjacent to a municipality or other urban area of more than 5,000 population (*1*). In Illinois rural areas, Interstate highway crashes totaled 6,298, and non-Interstate type crashes amounted to 509,514 (*1*).

This paper expands the traffic conflict concept in two principal ways: (*a*) freeway conflicts as opposed to intersection conflicts are examined and (*b*) freeway conflicts are counted from a microscopic freeway simulation program (which has been validated in its modeling of freeway traffic operations) rather than from direct empirical observations.

An overall goal of this work is to promote the safe movement of people and goods upon entering the freeway weaving section until they exit the weaving section. This goal can be expressed in two specific objectives: (*a*) to develop a tool that can be used to determine freeway conflict rates for an existing or proposed freeway facility and (*b*) to enhance the body of knowledge relating conflicts to crashes and the potential of using conflicts as an indicator of safety.

The first objective entailed the development of subroutines to be added to a freeway simulation program to assess the traffic conflicts that may occur in freeway weaving sections. The output of the revised program provides the user with a safety measure of the weaving section facility that is being modeled. In this way, highway engineers can examine the weaving traffic operations of a specific alternative as well as acquire an estimate of how safe the alternative will be if implemented.

METHODOLOGY

Definition of Terms

Simple Freeway Weaving Section

A simple freeway weaving section represents the physical space along a freeway where two traffic streams weave, as shown in Figure 1. The length of the weaving section is measured from a point where the entrance gore is 2 ft wide to a point where the exit gore is 12 ft wide, as shown in Figure 1 (*2*, pp. 4-1 to 4-19). This length represents the geometric distance in which weaving maneuvers must actually occur. In simple freeway weaving sections, there are four general movements: freeway to freeway, freeway to off-ramp, on-ramp to freeway, and on-ramp to off-ramp. Weaving traffic is composed of traffic from the second and third movements. The scope of this work is limited to ramp weaves because they are the most common type of freeway weaving sections.

Freeway Conflict

Two evident types of traffic conflicts occur on freeways. On the mainline, a driver is either following another vehicle or

J. Fazio, P/R Michaels Associates, Ltd., 3905 N. Long, Chicago, Ill. 60641. J. Holden, Research Development Office, Emergency Medicine, University of Illinois at Chicago, 1819 West Polk Street (M/C 724), Chicago, Ill. 60612–7354. N. M. Rouphail, Department of Civil Engineering, Mechanics, and Metallurgy (MC 246), University of Illinois at Chicago, P.O. Box 4348, Chicago, Ill. 60680.

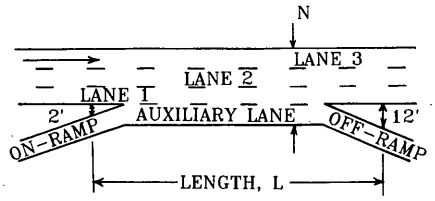


FIGURE 1 Layout of a ramp weave.

is in the process of changing lanes. Thus, the two most common types of freeway conflicts are rear-end conflicts and lane change conflicts. Some not so obvious conflicts are the head-on conflict, the object on lane of travel conflict, and the moving violation conflict. Their rare occurrence and lack of operational significance under normal driving conditions preclude any further discussion of their nature in this paper.

A freeway lane change conflict is a potential sideswipe or angle crash. It occurs when a vehicle changes lanes and the driver of the vehicle immediately following it in the target lane reacts to avoid a collision by applying the vehicle's brakes, as shown in Figure 2. If this maneuver is successful, the incident is a lane change conflict. Otherwise, a sideswipe or angle crash occurs.

Since the Integrated Traffic Simulation (INTRAS) program microscopically simulates traffic flow through its car-following and lane-changing algorithms, traffic conflicts occur in the simulations. However, in its original form, freeway conflicts were not counted in INTRAS, nor were they presented in its output reports. Thus, computer subroutines were developed, added to INTRAS, and microscopically validated so that freeway lane change and rear-end conflicts were accurately and precisely counted, their rates calculated, and conflict information reports generated as output. Moreover, the internal logic and validated car-following and lane-changing algorithms were not modified in any manner.

A lane change conflict was counted in the INTRAS simulation when the following two conditions were satisfied: the lane identification attribute of a freeway vehicle changed at a simulation time step (i.e., Vehicle A changed lanes in Figure 2) and the immediately following vehicle in the target lane (i.e., Vehicle B in Figure 2) had a deceleration greater than or equal to 0.61 m/sec^2 (2 ft/sec^2).

A freeway rear-end conflict is a potential rear-end crash. It occurs when a vehicle slows or stops on a freeway lane of travel, and the driver of the immediately following vehicle on the same freeway lane of travel reacts by applying the vehicle's brakes to avoid collision, as shown in Figure 3. If this maneuver is successful, the incident is a rear-end conflict. Otherwise, a rear-end collision occurs.

Counting rear-end conflicts in INTRAS was more difficult than counting lane change conflicts. In car-following, a vehicle

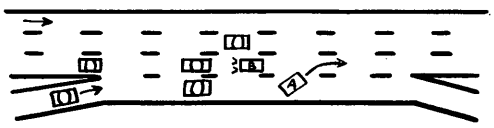


FIGURE 2 Lane change conflict: Vehicle A changes lane, and driver of Vehicle B brakes to avoid collision.

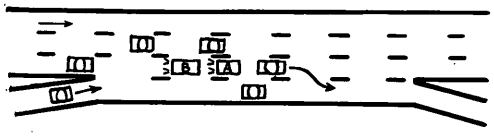


FIGURE 3 Rear-end conflict: Vehicle A slows or stops, and driver of Vehicle B brakes to avoid collision.

in a given lane may decelerate through one or more simulation time steps (i.e., it experiences a deceleration cycle). A rear-end conflict was counted in the INTRAS simulation when the following two conditions were satisfied: (*a*) Vehicle A in Figure 3 experiences a deceleration cycle and (*b*) during the deceleration cycle of Vehicle A, the immediately following vehicle (i.e, Vehicle B in Figure 3) experiences a full or partial deceleration cycle that has a maximum amplitude of 0.61 m/sec^2 (2 ft/sec^2) or greater. In other words, Vehicle A is decelerating while Vehicle B has a braking deceleration.

In an ideal situation, individual lane changes caused by weaving traffic do not result in lane change conflicts or in upstream rear-end or lane change conflicts. However, this ideal scenario is never sustained in real-world freeway traffic flow. Even on basic freeway segments, there is always a certain "background" lane-changing frequency that often results in some lane change conflicts. The additional impact of conflicts caused by weaving adversely affects traffic and increases speed variation in the weaving section. This vehicular speed variation essentially describes turbulence (i.e., major speed differences among vehicles in a segment of freeway). One measure of turbulence for weaving sections has been the difference between weaving and nonweaving vehicular speeds. These speeds were produced as output in the simulation runs. Thus, conflicts cause drivers to change their speeds. Reducing turbulence should enhance the safety of the system as well as its operations.

In a ramp weave, one of the more common type of weaving sections, turbulence is mostly concentrated in the auxiliary and rightmost freeway lanes; this is where most of the drivers perform their weaving maneuvers (i.e., interact with other entering or exiting drivers to reach desired destinations). By definition, weaving involves a certain amount of lane changing. At the microscopic level, this "crossing" inside ramp weave involves lane changes by drivers who want to complete their weaving maneuvers. Some of these lane changes are lane change conflicts; these incidents, and the additional conflicts they propagate, are the cause of speed variations at the macroscopic level (i.e., turbulence) and hazardous freeway operations.

To examine the association between freeway weaving section crash rates and conflict rates, crash rate and conflict rate data were obtained from existing freeway weaving sections. Crash rate data were obtained from the Illinois State Toll Highway Authority (ISTHA). The conflict rate data were determined by simulation modeling of the weaving sections where crash rates were obtained.

Sites

Site data were provided by ISTHA. Ten weaving section sites were selected as indicated in Table 1. The sites had histories

TABLE 1 Weaving Section Sites

Site	Location	Highway post marker	Length in kilometer (mile)
1.	Northbound I-294 (Tri-State Tollway) at Halsted Street (IL. 1)	2.77	1.1 (0.7)
2.	Southbound I-294 (Tri-State Tollway) at Halsted Street (IL. 1)	2.77	1.1 (0.7)
3.	Northbound I-294 (Tri-State Tollway) at Cicero Avenue (IL. 50)	12.08	0.6 (0.4)
4.	Northbound I-294 (Tri-State Tollway) at Ogden Avenue (U.S. 34)	27.63	1.1 (0.7)
5.	Southbound I-294 (Tri-State Tollway) at Ogden Avenue (U.S. 34)	27.63	1.1 (0.7)
6.	Northbound I-294 (Tri-State Tollway) from eastbound East-West Tollway (I-88) on-ramp to westbound Eisenhower Expressway (I-290) off-ramp	31.69	1.4 (0.9)
7.	Southbound I-294 (Tri-State Tollway) to eastbound Eisenhower Expressway (I-290) off-ramp from westbound East-West Tollway (I-88) on-ramp	31.69	1.4 (0.9)
8.	Northbound I-294 (Tri-State Tollway) from eastbound Kennedy (I-190) on-ramp to westbound Northwest Tollway (I-90) off-ramp	40.33	1.3 (0.8)
9.	Northbound I-294 (Tri-State Tollway) at Grand Avenue (IL. 132)	69.61	1.3 (0.8)
10.	Southbound I-294 (Tri-State Tollway) at Grand Avenue (IL. 132)	69.61	1.3 (0.8)

of high crash occurrence. All the sites are located on Interstate 294, the Tri-State Tollway. The Tri-State Tollway is a major cross-country freight route and also serves as a quasi-beltway for the Chicago metropolitan area.

Determination of Crash Rates

Crash counts by type (i.e., rear-end and angle/sideswipe) and by year were extracted from ISTHA computer printouts. The printouts contained only those crashes in or near the 10 weaving sections as indicated by the milepost of each crash record. The crash counts were tallied by year for the period 1985 through 1988. The crashes could have been disaggregated by a shorter interval of time, but such a strategy would increase crash count variability. Some of the crash counts at certain sites had to be aggregated with another site. This occurred when the sites had the same mileposts but opposite direction of traffic. The printouts did not distinguish whether the direction was northbound or southbound. Thus, the crash counts for the two Halsted Street sites, the two Ogden Avenue sites, the two I-88/I-290 sites, and the two Grand Avenue sites are combined.

To convert the crash counts to rates, the directional annual average daily traffic (AADT) through the weaving section and the length of the segment on which the crashes occurred must be determined. The directional AADT schematics of the 10 weaving sections for January 1, 1985, through July 31, 1988, were obtained from ISTHA. The AADT through the weaving section was known, and AADT for the on-ramp, off-ramp, freeway upstream, and freeway downstream of the section were also given. This additional information was used in determining origin and destination traffic movements when the weaving sections were modeled for simulation. Mainline AADTs and some ramp AADTs are determined from continuous counts throughout the year because of the mainline tollway plazas and ramp toll machines on the Tri-State. AADTs of ramps that had no collection devices were based on daily counts taken for a week and seasonally adjusted. Only one ramp had a toll collection device in the 10 sites examined—the on-ramp to Site 2. In general, AADT increases with each subsequent year except at Site 1 in 1988. In 1987 and 1988, this site was reconstructed into a collector-distributor weaving section.

The highway mainline lengths where the crashes occurred were determined from the upper and lower milepost values from the middle of the site. This milepost range is determined from the crash printouts. Once the distance of this range is known, 161 m (0.1 mi) are added [i.e., 80.5 m (0.05 mi) for each end] to adjust for round-off error in the crash reporting. Table 1 gives the length at each site where the crashes were counted.

Once the crash counts, AADT, and weaving section length were determined, the crash rates were calculated by crash type and year, as given in Table 2 and shown in Equations 1 and 2 for 1985 through 1987 and 1988, respectively:

$$\text{crash rate} = \text{crash count}/(\text{AADT} * \text{length} * 365) \quad (1)$$

$$\text{crash rate} = \text{crash count}/(\text{AADT} * \text{length} * 213) \quad (2)$$

TABLE 2 Weaving Section Crash Rates (Crashes per 100 Million Vehicle Kilometers)

Site	Year	Crash Type				
		Rear-end	Sideswipe + Angle	Overturn	Other	Total
1. and 2.	1985	63.4 (102)[b]	79.5 (128)	0 (0)	43.2 (69.5)	186.4 (300)
	1986	26.6 (42.8)	35.5 (57.1)	0 (0)	20.7 (33.3)	82.6 (133)
	1987	16.2 (26.0)	48.5 (78.1)	0 (0)	18.9 (30.4)	83.9 (135)
	1988[a]	36.2 (58.3)	65.2 (105)	0 (0)	65.2 (105)	166.5 (268)
3.	1985	114.3 (184)	157.2 (253)	0 (0)	142.9 (230)	413.8 (666)
	1986	73.9 (119)	98.8 (159)	0 (0)	61.8 (99.4)	234.9 (378)
	1987	33.6 (54.0)	22.4 (36.0)	0 (0)	11.2 (18.0)	67.1 (108)
	1988[a]	0 (0)	17.1 (27.5)	0 (0)	17.1 (27.5)	34.2 (55.0)
4. and 5.	1985	22.9 (36.8)	16.0 (25.8)	0 (0)	9.1 (14.7)	48.0 (77.3)
	1986	27.3 (43.9)	25.2 (40.5)	0 (0)	6.3 (10.1)	58.8 (94.6)
	1987	17.6 (28.4)	25.5 (41.0)	0 (0)	9.8 (15.8)	52.9 (85.1)
	1988[a]	40.6 (65.3)	0 (0)	0 (0)	12.5 (20.1)	53.1 (85.4)
6. and 7.	1985	57.6 (86.2)	60.3 (97.0)	0 (0)	21.7 (35.0)	135.5 (218)
	1986	72.7 (117)	56.4 (90.8)	0 (0)	6.5 (10.4)	135.5 (218)
	1987	111.2 (179)	63.4 (102)	0 (0)	15.0 (24.2)	189.5 (305)
	1988[a]	149.1 (240)	81.4 (131)	0 (0)	12.0 (19.3)	242.3 (390)
8.	1985	94.4 (152)	113.1 (182)	9.5 (15.2)	51.8 (83.4)	268.4 (432)
	1986	131.1 (211)	103.8 (167)	4.5 (7.26)	40.6 (65.4)	279.6 (450)
	1987	119.3 (192)	80.8 (130)	0 (0)	51.2 (82.4)	251.7 (405)
	1988[a]	154.7 (249)	60.5 (97.3)	0 (0)	26.8 (43.2)	241.7 (389)
9. and 10.	1985	16.2 (26.1)	10.8 (17.4)	0 (0)	16.2 (26.1)	43.2 (69.6)
	1986	15.3 (24.6)	15.3 (24.6)	0 (0)	10.2 (16.4)	40.7 (65.5)
	1987	30.4 (49.0)	25.4 (40.8)	0 (0)	45.7 (73.5)	101.3 (163)
	1988[a]	76.4 (123)	15.3 (24.7)	0 (0)	69.0 (111)	160.9 (259)

[a]From January 1, 1988 to July 31, 1988, i.e., 213 days
[b]Crashes per 100 million vehicle-miles

Determination of Conflict Rates

The INTRAS program, which is distributed by FHWA under its technology transfer program, was implemented to model the 10 existing ramp weaves. Certain subroutines were added to the INTRAS program to allow the extraction of conflict data in the ramp weaves. The accuracy of the conflict counts from the simulations was established by examining vehicle trajectory dumps and the "flags" that mark the time step when a conflict was counted. In this way, the simulation conflict counting algorithms were debugged until they precisely and accurately matched the definitions of the rear-end and lane change conflict previously mentioned. Once the conflict counts were determined for a specific time period, they were divided by the vehicle-kilometers (vehicle-miles) of travel in the weaving section.

There are several reasons why INTRAS was selected for adaptation in this paper. First, to count individual vehicle conflicts, a microscopic simulation program is needed; INTRAS is microscopic. The fact that INTRAS has been rigorously validated at weaving sections gives credibility to its results (*3,4*). The capability of INTRAS to control for geometric and volume variables important to weaving sections (e.g., configuration, freeway volume, ramp volumes, length, and number of lanes) is another reason for selection. Next is an economic consideration: using INTRAS to generate and collect weaving section data is much less costly than collecting and processing data from the field. Finally, and most important, INTRAS uses highly detailed lane change and car-following logic. Such an elaborate simulation provided much needed insight and understanding of the complex turbulence relationships in a weaving section.

Once the 10 sites were modeled into INTRAS, traffic flow through each site was simulated for 15 min after equilibrium was obtained. Two hundred computer runs were performed (10 sites × 4 years × 5 random number seeds). In each computer run, total lane change conflict rates, mandatory lane change conflict rates, and rear-end conflict rates were recorded. The rates were determined by examining the appropriate INTRAS output reports and freeway link. One of the drawbacks of conflict rates, like crash rates, is that they can be highly variable. Yet, the variations in conflict occurrences by changing the random number seed in the simulation should not be large, that is, coefficient of variation (CV) ≤ 1, given all else held constant because of the stochastic nature of vehicle generation. By varying the random number seed five times, the mean, standard deviation, and CV for each group of five resulting conflict rates were also calculated. The CV values were quite low in all three conflict rate categories; most values were less than 1. The only exceptions occurred when the conflict counts were very low; this was expected.

Statistical Analyses

The Spearman rank correlation procedure was used to test the correlation between total lane change conflict rates and angle/sideswipe crash rates and the correlation between rear-end conflict rates and reported rear-end crash rates. The rank correlation coefficient of Spearman (r_S) is used in

general when a population does not have an approximate bivariate normal distribution; the test is exact for small sample sizes and nonnormal data, and the effects of outliers are weakened (*5*).

The Spearman rank-correlation procedure has been in use since 1904; its main advantages are simplicity and power. In addition, the method has been proven to be almost as powerful as its classical counterpart, the Pearson "product-moment" correlation method, under conditions favorable to the latter method and even more powerful than the parametric method when the Spearman method's assumptions are violated (*6*).

The Spearman rank-correlation procedure consists of five steps:

1. Replace the n values of the conflict rates by their ranks by giving the rank 1 to the "largest" and the rank n to the "smallest" (note: no ties were observed in the data).
2. Replace the n values of the crash rates by their ranks as in Step 1.
3. For each of the n rates, obtain a set of rank difference scores, that is,

$$D_i = \text{conflict rate rank}_i - \text{crash rate rank}_i$$

where $i = 1, 2, \ldots, n$.

4. Obtain the summation of squared rank difference scores, ΣD_i^2.
5. Obtain the Spearman rank-correlation coefficient (r_S), where

$$r_S = 1 - \{(6\Sigma D_i^2)/[n(n^2 - 1)]\} \quad (3)$$

The r_S ranges from -1 to $+1$. When r_S is -1, a perfect negative relationship exists. The two variables are independent of each other when $r_S = 0$. When $r_S = +1$, a perfect positive relationship exists. After r_S has been determined, a hypothesis test can be performed using the proper test statistic (*5*).

RESULTS

Before any test could be performed, three cases had to be discarded. Two cases were dropped because the ramp weave was reconstructed into a collector-distributor weaving section. The crash rates of these two cases would be biased because of the construction activity. The third case was ignored because its high volume levels could not be maintained by the simulation program. Thus, in the revised data base, there was a total of 21 observations.

Hypothesis testing was conducted using a $p = 0.025$ level of significance. The null hypothesis is stated as

H_o: Crash rates and conflict rates are independent.

The alternative hypothesis is

H_A: Low crash rates and low conflict rates tend to occur together, as do high crash rates and high conflict rates. In summary, crash rates and conflict rates are positively correlated.

The decision rule for the one-sided test at a 0.025 level of significance is

$$\text{Reject } H_o \text{ if } |r_S| \geq \text{critical value } r_{S,n,0.025}$$

where $|r_S|$ is the absolute value of the Spearman's rho correlation coefficient,

$$|r_S| = |1 - \{[6 * \Sigma_1^n (\text{rank difference}_i)^2]/[n * (n^2 - 1)]\}| \quad (4)$$

and $r_{S,n,0.025}$ is the test correlation coefficient obtained from an appropriate table (*5*).

The entire data set was disaggregated by three weaving section length ranges: long, moderate, and short. The long categories involved those three cases with lengths between 594.4 m (1,960 ft) and 624.8 m (2,050 ft). The eight cases with lengths between 259.1 m (850 ft) and 304.8 m (1,000 ft) were defined as having moderate length. The 10 cases with lengths between 152.4 m (500 ft) and 198.1 m (650 ft) were defined as short. Spearman tests were attempted for each length of the three categories for the lane change rates, as presented in Table 3, and similarly three Spearman tests were attempted for the rear-end rates, as presented in Table 4.

For the long length cases involving rear-end and lane change rates, valid Spearman's tests could not be conducted because of the small sample size of three cases. The minimum sample size required to obtain a test statistic is four (*5*). A desirable sample size would be greater than four in order to perform hypothesis tests on the significance of Spearman correlation coefficients. Although hypothesis tests could not be conducted on the long length cases, moderate Spearman correlation coefficients of +0.50 were calculated for the lane change and rear-end rates. The four CV values of the total lane change conflict rate, angle/sideswipe crash rate, rear-end conflict rate, and rear-end crash rate data were extremely low. The low values imply that the conflict rates and the crash rates did not vary widely in cases where the weaving sections had a long length.

For the eight moderate length cases, the results of the Spearman tests indicate that the null hypothesis is rejected at the $p = 0.025$ level of significance for the lane change rate test and for the rear-end test, and the null hypotheses were also rejected at $p = 0.050$. Thus, one can be 97.5 percent confident that there is a positive correlation between total lane change conflict rates and reported angle/sideswipe crash rates and between rear-end conflict rates and reported rear-end crash rates.

For the 10 short length cases, the two Spearman test results indicate that their null hypotheses could not be rejected at a level of significance of $p = 0.025$ for either the lane change rate test or the rear-end rate test, but at the $p = 0.050$ level they were rejected. Thus, at short weaving section lengths, no significant correlation existed between the conflict rates and crash rates at $p = 0.025$. The CV values were consistently equal or higher than their CV value counterparts in other length categories. Furthermore, the correlations were negative; this indicates an inverse relationship between conflict and crashes at the short length sites.

There are many reasons for the high variations in the conflict rate and crash rate data especially when the weaving

TABLE 3 Spearman Test Results by Length for Lane Change Rates

Site Number	Number of Ranks	TOTAL LANE CHANGE CONFLICT RATE Mean[d]	CV	SIDESWIPE AND ANGLE CRASH RATE Mean[e]	CV	Spearman Correlation Coefficient
6 and 7[a]	3	0.1251 (0.2014)[i]	0.0557	60.0 (96.5)[j]	0.0562	+0.50[f]
8, 9 and 10[b]	8	0.0337 (0.0542)[i]	0.6004	53.1 (85.5)[j]	0.7927	+0.74[g]
1 and 2, 3, 4 and 5[c]	10	0.0747 (0.1202)[i]	0.5977	47.7 (76.8)[j]	1.0264	-0.61[h]

[a]Long Length Weaving Sections, 594.4 m (1,950 ft) to 624.8 m (2,050 ft)
[b]Moderate Length Weaving Sections, 259.1 m (850 ft) to 304.8 m (1,000 ft)
[c]Short Length Weaving Sections, 152.4 m (500 ft) to 198.1 m (650 ft)
[d]In conflicts per vehicle-kilometers (cpvkm)
[e]In crashes per 100 million vehicle-kilometer
[f]Insufficient sample size for hypothesis test
[g]Significant at 0.025 and 0.050
[h]Not significant at 0.025, significant at 0.050
[i]In conflicts per vehicle-mile
[j]In crashes per 100 million vehicle-mile

sections were of a short length. Errors in crash reporting, errors in determining AADT, errors between the lengths when the crashes were counted and weaving section length, and the stochastic nature of the simulation program probably contributed to data variation.

Two general conclusions were noted in all Spearman tests when the data were disaggregated regardless of whether the test involved lane change or rear-end rates. One is that when there is a decrease in the crash rate CV, there usually is a corresponding decrease in its conflict rate CV counterpart. Likewise, when the crash rate CV increases, usually so does the conflict rate CV. The second conclusion is that when the magnitude of the crash rate CV values are high, their conflict rate CV value counterparts tend to be high. When the crash rate CVs were low, the conflict rate CVs tend to be low. These two conclusions shed an important light on the freeway conflict-crash relationship: freeway weaving section conflict rate variations mimic the directional changes and magnitude of the freeway weaving section crash rate variations.

Another interesting finding is that the mean crash rates by length partially confirm previous research regarding crash rate versus weaving section length (*7*). Crash rates tend to stabilize for ramp weaves with lengths greater than 228.6 m (750 ft). When average rear-end and lane change crash rates are added together by long and moderate length categories that are greater than 228.6 m (750 ft), the average crash rates were 133.0 crashes per 100 million vehicle-kilometers (mvkm) [214 crashes per 100 million vehicle-miles (mvm)] for the moderate length category and 139.2 crashes per 100 mvkm (224 crashes per 100 mvm) for the long length category. The short length category had a crash rate of 89.5 crashes per 100 mvkm (144 crashes per 100 mvm). There was only a 4.7 percent difference between long and moderate length rates, whereas a 33.7 percent difference existed between short and moderate length rates.

TABLE 4 Spearman Test Results by Length for Rear-End Rates

Site Number	Number of Ranks	REAR-END CONFLICT RATE Mean[d]	CV	REAR-END CRASH RATE Mean[e]	CV	Spearman Correlation Coefficient
6 and 7[a]	3	2.7333 (4.3988)[i]	0.1944	78.9 (127)[j]	0.3715	+0.50[f]
8, 9 and 10[b]	8	0.5130 (0.8256)[i]	0.5291	79.5 (128)[j]	0.6810	+0.95[g]
1 and 2, 3, 4 and 5[c]	10	1.7009 (2.7373)[i]	0.7287	42.0 (67.6)[j]	0.7911	-0.48[h]

[a]Long Length Weaving Sections, 594.4 m (1,950 ft) to 624.8 m (2,050 ft)
[b]Moderate Length Weaving Sections, 259.1 m (850 ft) to 304.8 m (1,000 ft)
[c]Short Length Weaving Sections, 152.4 m (500 ft) to 198.1 m (650 ft)
[d]In conflicts per vehicle-kilometers (cpvkm)
[e]In crashes per 100 million vehicle-kilometer
[f]Insufficient sample size for hypothesis test
[g]Significant at 0.025 and 0.050
[h]Not significant at 0.025, significant at 0.050
[i]In conflicts per vehicle-mile
[j]In crashes per 100 million vehicle-mile

In summary, the results demonstrate a positive correlation between lane change conflict rates and reported sideswipe/angle crash rates and between rear-end conflict rates and reported rear-end crash rates at ramp weaves of moderate length. The conflict rate CV (i.e., standard deviation divided by the mean) is proportional to the crash rate CV; as the crash rate CV changes so does the conflict rate CV in the same direction. Furthermore, when the crash rate CV has a high value, the conflict rate CV tends to be high, and vice versa. The total rear-end plus lane change crash rate stabilizes when the weaving section length is greater than 259.1 m (850 ft).

Another observation from the results was that the average conflict rates for the weaving sections with short lengths are higher than the conflict rates of the moderate length sections. Average rear-end conflict rate for the short sections was 232 percent higher than the moderate sections. The lane change rate was 122 percent higher. Thus, weaving sections with short lengths have higher rear-end conflict rates and lane change conflict rates than sections of moderate length given similar four volume movements. These higher conflict rates translate into more traffic turbulence in those weaving sections.

The average conflicts rates in the long length category were abnormally high. The two sites in this category (northbound and southbound Tri-State near I-88) are well-known problem sites according to ISTHA officials. In fact, plans are being formulated for the reconstruction of these two sites. Besides length, they are different from most of the other sites in the data base because of their high nonweaving and weaving volumes and because their on- and off-ramps connect to other Interstate highways (I-88 and I-290), not to arterials like the moderate and short length sites. The higher volumes translated into a higher probability for a conflict to occur, despite the long length. The ability of the model to identify weaving sections with high conflict rates, as exemplified in the preceding two cases, confirms that the model can be used by the highway engineer to identify weaving sections that are hazardous to the motoring population. In other words, conflict rates of Interstate highway ramp weaves can be used as an indicator of traffic safety in Interstate highway ramp weaves.

Average rear-end and lane change crash rates were actually less in the weaving sections with short lengths than the rates in the other length categories. This observation does not confirm previous research indicating that crash rates in weaving sections increase significantly as weaving section length decreases after 228.6 m (750 ft), as indicated in Table 5 (*7*). This observation implies that either the average crash rates for moderate or long length sites were excessively high or the crash rate was abnormally low for the short length category, or both. At first, this discrepancy was thought to be directly due to difference between the length in which the crashes were counted and the measured weaving section length. However, a check on the differences in these lengths indicated that the difference was constant at approximately 914.4 m (3,000 ft) for 9 of the 10 sites. It was then thought that errors in the method by which AADT was determined might cause a bias in the crash rates. If there was a bias, there is no reason to believe that it would be more so for a specific length category (i.e., the bias would affect all sites equally). The only remaining explanation for the discrepancy lies in the process in which the crash locations were reported. Apparently, when the weaving section had a moderate to long length, more crashes tended to be reported in the segment that contained the weaving section. Similarly, fewer crashes tended to be attributed to the segments that contained weaving sections of short lengths.

An implicit assumption in the data base disaggregated by length category is that the yearly observations of crash rates for a specific site are independent of each other. To verify this assumption, an S_3 sign test of Cox and Stuart for the detection of a monotonic trend was performed using the total crash rate data in Table 2. The S_3 sign test was applied to determine whether the time series data (yearly crash rates) had no trend (H_o) (i.e., were random) or had a monotonic trend (H_A). The z-statistic equation used was one for small samples ($n < 30$): $z = [\mathrm{ABS}(S - n/6) - 0.5]/[\mathrm{SQRT}(n/12)]$, where n is the number of observations rounded to an appropriate multiple of 3 and S is the sum of the plus or minus signs (*5*). The results of the S_3 sign test indicated that an increasing or decreasing trend in the crash rates involving each

TABLE 5 Comparison of ISTHA Crash Rates with Past Research

n[a]	Length in meters (feet)	Mean VF[b] (vph)	Mean VRP[c] (vph)	Mean Crash Rate[d] Per 100 mvkm (Per 100 mvm)	Mean Crash Rate[e] Per 100 mvkm (Per 100 mvm)
3	594.4 to 624.8 (1950 to 2050)	3852	1674	139.2 [285.8][f] (224)	130.5 (210)
8	259.1 to 304.8 (850 to 1000)	2203	551	133.0 [54.7][f] (214)	130.5 (210)
10	152.4 to 198.1 (500 to 650)	3996	360	89.5 [177.7][f] (144)	273.4 to 155.3 (440 to 250)

[a] ISTHA site observations

[b] VF is the entering freeway volume to the ramp weave in vehicles per hour

[c] VRP is the entering ramp volume to the ramp weave in vehicles per hour

[d] ISTHA, Angle/sideswipe and rear-end crash rates

[e] Reference (*7*)

[f] Value in brackets indicates rear-end plus total lane change conflict rates in 0.01 conflicts per vehicle-kilometers

of the different sites or combination of sites could not be ascertained at the $p = 0.050$ level.

DISCUSSION OF RESULTS

Crash rates do provide the highway engineer with a measure by which to evaluate the level of safety of a facility. However, they are generally not applied in the determination of the quality of traffic flow through the highway facility. Implicitly, freeway conflicts have a disruptive operational effect on freeway traffic flow; they are a source of turbulence. Conflict rates can be used as a measure of the level of service (LOS) of traffic operations in ramp weaves because they quantify freeway turbulence, especially when the freeway facility has moderate to high service flows. Currently, average weaving and nonweaving speeds of vehicles through the ramp weave are applied in operational analysis procedures to determine the LOS of weaving and nonweaving traffic (*2*). In regard to freeway ramp weaves, conflict rates are minimized when the traffic through them operates at LOS C or better (*8*). Thus, when designing ramp weaves to handle forecast peak traffic flow rates, a geometry should be selected such that the weaving and nonweaving traffic operates at LOS C or better.

Conflicts may also be analyzed to determine the level of safety of a freeway facility. Conflicts do not have to be associated with crashes to be a good indicator of safety. However, it is desirable to examine conflict/crash associations to strengthen the argument that conflicts may be used as a surrogate for crashes in lieu of crash data. A good indicator of traffic safety is characterized by the completeness of the counts on which the indicator is dependent, the frequency of occurrence (sufficient sample size), and its ability to identify hazardous facilities and prevent or minimize hazardous events and situations.

In comparing the advantages and disadvantages of using either conflict rates or crash rates as an indicator of safety, crash rates have some major disadvantages. First, not all crashes are reported, especially the very minor ones. A crash rate based on such counts may not be representative of the crash population. Second, crashes are usually reported by law enforcement officials and transcribed into a data base by a data entry operator. This process of reporting and recording crashes could lead to ambiguities relative to the exact location and time of crash occurrence and transcription errors. Both may result in misleading crash statistics for a given facility. This could have been why negative correlations between conflicts and crashes were observed for the short length sites; reporting personnel may have underreported crashes in short weaving sections by using mileposts numbers outside the ramp weaves. Third, most crash rates use vehicle-kilometers (vehicle-miles) of travel as an indicator of risk. Using vehicle-kilometers (vehicle-miles) as an exposure measure is better than using raw crash counts, but it has serious drawbacks. This exposure measure does not consider the total number of passengers in the vehicle, speed (i.e., a time element is not involved), or the accuracy of the procedures from which the number of vehicles is obtained (e.g., AADT and gasoline mileage).

Conflict rates should be a good indicator of traffic safety because they are predictors of the probability of crash occurrence. Thus, one would not need to wait for crash occurrence to assess the level of safety of a traffic facility. Conflict rates are easily measured from the field over a short time and distance interval. For a ramp weave, conflict rate determination using empirical methods involves counting vehicular brake light indications over a peak 15-min interval along the weaving section length and the total number of vehicles that entered the ramp weave. The counts can be electronically stored at the time of collection and directly uploaded into a computer for data processing. These counts would not have the ambiguities and errors associated with crash reporting and processing. The only major source of error would be malfunction of vehicular brake lights. Given that newer vehicles have three brake lights (right, left, and center rear window), the probability of all three being simultaneously inoperable is low. Finally, the work reported here indicates that conflict rates should mimic crash rates in terms of variation (i.e., when crash rate variation increases, conflict rate variation also increases, and vice versa).

The factors that contribute to crash occurrence on the freeway mainline also contribute to conflict occurrence. Such factors are the driver, freeway geometry, the vehicle, and the environment. For drivers, items such as reaction time, vision, age, experience, and blood alcohol level may contribute to crash/conflict occurrence. Vehicle characteristics such as type and condition (brake malfunctioning, excessive tire wear, etc.) may be contributing factors. Freeway geometries such as sharp horizontal curves, steep grades, short weaving section length, and short speed change lanes are items that may affect crash/conflict occurrence. Environmental factors that may be important are day-to-day weather and road surface conditions, for example, fog, glare, and ice on the roadway. If crashes are aggregated over a yearly basis, the temporary effect of the weather will average out over the years.

CONCLUDING REMARKS

A Spearman's correlation coefficient of $+0.74$ occurred between lane change conflict rates and reported angle/sideswipe crash rates, and a coefficient of $+0.95$ occurred between rear-end conflict rates and rear-end crash rates, for the eight ramp weaves that had moderate lengths between 260 m (850 ft) and 305 m (1,000 ft). Both these correlations were significant at $p = 0.025$ and $p = 0.050$. For the 10 weaving sections with short lengths between 152 m (500 ft) and 198 m (650 ft), the lane change and rear-end hypothesis tests indicated no significant correlations between the conflict and crash rates at $p = 0.025$, but significance at $p = 0.050$. The two long sites with lengths of 610 m (2,000 ft) had moderate Spearman's correlation coefficients of $+0.50$ between lane change conflict rates–angle/sideswipe crash rates and between rear-end conflict rates and rear-end crash rates; an insufficient sample size prevented hypothesis tests from being conducted.

With the developed conflict extracting subroutines that were added to FHWA's freeway simulation program (INTRAS, soon to be renamed FRESIM), conflict output reports are produced to enable the highway engineer to obtain not only operational performance measures of a freeway facility but also a measure of safety. Properly using the program, the engineer may model proposed freeway facility alternatives to determine which design would be less hazardous. In this pa-

per, only weaving section facilities were examined. Moreover, conflict output reports are also produced for basic freeway segments and freeway ramp junctions. Further conflict/crash studies involving those two freeway components are recommended.

ACKNOWLEDGMENT

The authors want to thank Melvin Sierakowski and Adolf Green of the Illinois State Toll Highway Authority for their assistance in identifying the top 10 hazardous weaving sections on the tollway system and providing crash data from their computer data base.

REFERENCES

1. *Accident Facts 1988.* Illinois Department of Transportation, Springfield, 1988.
2. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985.
3. A. Skabardonis, M. Cassidy, A. D. May, and S. Cohen. The Application of Simulation To Evaluate the Operation of Major Freeway Weaving Sections. In *Transportation Research Record 1225,* TRB, National Research Council, Washington, D.C., 1989, pp. 91–98.
4. R. B. Goldblatt. *Development and Testing of INTRAS, a Microscopic Simulation Model, Vol. 3, Validation and Application.* Report FHWA/RD-80/108. FHWA, U.S. Department of Transportation, 1980.
5. L. Sachs. *Applied Statistics, A Handbook of Techniques* (5th edition). Springer-Verlag New York, Inc., 1982.
6. M. L. Berenson and D. Levine. *Basic Business Statistics, Concepts and Applications* (3rd edition). Prentice-Hall, Englewood Cliffs, N.J., 1986.
7. J. C. Fee. Accident Experience on Speed-Change Lanes of the Interstate Highway System. *Public Roads,* Vol. 37, 1972, pp. 61–64.
8. J. Fazio. *Modeling Safety and Traffic Operations at Freeway Weaving Sections.* Ph.D. dissertation. University of Illinois at Chicago, Chicago, 1990.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Association of Median Width and Highway Accident Rates

MATTHEW W. KNUIMAN, FORREST M. COUNCIL, AND DONALD W. REINFURT

Data for two states have been extracted from the Highway Safety Information System and used to examine the effect of median width on the frequency and severity of accidents. Log-linear models for accident rates have been used to describe the effect of median width after adjusting for other variables. Effects have been estimated by the quasi-likelihood technique assuming a negative-binomial variance for the accident count per roadway section. Results for both states indicate that total accident rates and rates for specific types and severity decline rapidly when median width exceeds about 25 ft (7.6 m). Policy guidelines for median widths are somewhat nebulous, partly due to the lack of large well-conducted studies providing quantitative information on this topic. The results provide a basis for the development of more precise guidelines regarding median width.

Medians on divided highways provide a recovery area for out-of-control vehicles. The median should be wide enough to allow an out-of-control vehicle sufficient space to recover without crossing over the median into opposing traffic. In addition, divided highways with wide medians provide a safety zone at access points for turning vehicles and entering vehicles wishing to cross one or both directions of traffic. A variety of median types are in use, with narrow medians sometimes including barriers designed to positively prevent out-of-control vehicles from crossing the median into opposing traffic.

It has been suggested that the median width should be at least 60 ft (18.3 m) on rural highways and can be as low as 10 ft (3.1 m) on urban highways if median barriers are provided (*1*), but little research has been conducted providing quantitative measures of the effects of median width on the frequency and severity of related accidents. Early studies (*2–5*) were not able to establish definitive relationships between accident rates and median width; however, a subsequent study by Garner and Deen (*6*) has shown that wider medians have lower accident rates. The Garner and Deen study used 420 mi (676 km) of rural, four-lane, fully controlled access road sections [speed limit 70 mph (113 kph)] in Kentucky with median widths ranging from 20 to 60 ft (6.1 to 18.3 m) and involved a total of 2,448 accidents (1965–1968).

This paper examines the effect of median width on the frequency and severity of accidents on homogeneous highway sections with a traversable (nonbarrier) median. Highway sections with curbed medians or medians including barriers were also examined. However, there were insufficient sections of these types for meaningful statistical analysis.

Data extracted from the Highway Safety Information System (HSIS) for the states of Utah and Illinois are used. The Utah data involve 982 sections of highway for a total of 973.8 mi (1567.8 km) of roadway with 37,544 reported accidents over the period 1987 through 1990. The Illinois data involve 2,481 sections of highway for a total of 2,081.3 mi (3351 km) of roadway and 55,706 accidents over the period 1987 through 1989. Road sections with median widths ranging from zero (no median) to 110 ft (33.6 m) are examined.

METHODS

Data Base

HSIS developed and maintained for the Federal Highway Administration by the Highway Safety Research Center (HSRC) at the University of North Carolina, includes an accident data base, a road inventory data base, and a traffic volume file for five states (Illinois, Utah, Michigan, Minnesota, and Maine). All accidents reported to the police are included in the data base, and for each accident a variety of details are recorded, including date and location of accident, road and environmental conditions, accident type, and the number and severity of injuries. The road inventory data base contains the characteristics of homogeneous highway sections. The definition of homogeneous varies to some degree from state to state, but in most cases a new section is initiated any time there is a change in a major geometric or cross-section variable (e.g., lane width, pavement type, shoulder width or type, number of lanes, etc.). For this study, homogeneous sections of highway were defined as contiguous segments for which the following variables did not change: federal aid system, functional classification, rural/urban designation, predominant terrain type, average annual daily traffic volume (both directions), one- or two-way operation, number of lanes, average through lane width, posted speed limit, access control, median width and type, left shoulder width, and right shoulder type.

The traffic volume file contains the average annual daily traffic volume. Using route number and mile points, these three files can be merged to obtain the number, rate, severity, and type of accidents that have occurred on specific highway sections over a given period of time.

Extensive checking and preliminary investigation indicated that the accident and roadway data for two of the five states

M. W. Knuiman, Department of Public Health, University of Western Australia, M Block, QE II Medical Centre, Nedlands, Western Australia 6009, Australia. F. M. Council and D. W. Reinfurt, Highway Safety Research Center, University of North Carolina, CB3430, 134½ East Franklin Street, Chapel Hill, N.C. 27599.

(Utah and Illinois) were of adequate completeness and reliability for an analysis investigating the effect of median width on accident rates. The Utah and Illinois data were described by Council and Hamilton (*7*) and Council and Williams (*8*), respectively.

Several roadway characteristics in addition to median width affect the frequency, severity, and type of accidents. To isolate the effect of median width, these other variables must be controlled either by restricting the road sections to have particular characteristics or through statistical adjustment. In this study both methods of control were used.

The analyses have for the most part been restricted to two-way, four-lane, rural and urban Interstate, freeway, and major highway road sections of length exceeding 0.07 mi (0.11 km), with posted speed limit at least 35 mph (56 kph) and with median widths ranging from zero (no median) to 110 ft (33.6 m). A section length of 0.07 mi (0.11 km) was chosen as the minimum length for which reported accident locations could be considered reliable for merging with the road inventory data base. Sections on minor roads were eliminated because many had missing data and virtually all had no median. After these were eliminated, there were very few sections with speed limit less than 35 mph (56 kph), so the remaining few were also eliminated. There were also a few sections with median width ranging from 111 ft (33.9 m) to 999 ft (304.7 m), and these were eliminated because they were possibly in error and would have a large influence on the median width coefficients in a regression model. In addition, the Utah analysis was restricted to road sections with lane width of 12 ft (3.7 m). There was no explicit lane width variable for Illinois, and it could not be reliably calculated from other variables; thus no such restriction was applied for Illinois.

Median width is defined as the width of the portion of divided highway separating the traveled ways for traffic in opposite directions (and includes the inside shoulder). Other variables considered in the statistical analyses were as follows: functional classification (categorized as rural–Interstate/freeway, rural–other major road, urban–interstate/freeway, urban–other major road), posted speed limit [35 to 40, 45 to 50, 55, and 65 mph (56 to 64, 72 to 81, 89, and 105 kph)], right shoulder width, access control (full, partial, none), curvature (value 1 if curvature greater than 1 degree, 0 otherwise), average daily traffic (average number of vehicles per day), and section length [in miles (kilometers)]. Access control data were not reliable for Utah (on the basis of information from state data experts) and were therefore not considered in the Utah analysis. Furthermore, 23 percent of the Utah sections did not have speed limit recorded and thus an additional category "missing" was used for this variable. Curvature was not considered in the Illinois analysis because the data were incomplete.

The Utah analysis was based on 982 sections of highway for a total of 973.8 mi (1567.8 km) of roadway [average section length 0.99 mi (1.6 km)], and the Illinois analysis involved 2,481 sections of highway for a total of 2,081.3 mi (3350.9 km) of roadway [average section length 0.84 mi (1.35 km)].

For each Utah road section, the number of accidents over the 4-year period 1987–1990 was obtained (giving a total of 37,544 accidents), whereas for Illinois the 3-year period 1987–1989 was used (giving a total of 55,706 accidents). The 1990 Illinois data did not yet exist in the HSIS files at the time of this analysis. Each accident had a severity code representing the most serious injury in the accident (K = fatal, A = incapacitating injury, B = nonincapacitating injury, C = possible injury, PDO = property damage only). The number of total accidents and the number of each severity type were determined for each section of road for use in total, A + K, C + B + A + K (i.e., all injury), and PDO crash rates.

The accident data from both states also provided numerous variables concerning the nature of the accident, including accident type, collision sequence (in Utah), and vehicle movements preceding and during the accident sequence. An attempt was made to define a smaller number of accident categories based on "potential median involvement"—the degree to which the presence and width of a median might potentially affect the crash rate. This categorization was based on the assumption that the basic goals of a median are (*a*) to separate opposing vehicles, (*b*) to provide a vehicle with a safe clearzone that can be used to avoid vehicles traveling in the same direction, (*c*) to provide a refuge for turning or crossing vehicles, and (*d*) to provide a safe clearzone to reduce the number of ran-off-road object impacts. In the resulting categorization, each accident was coded as a multivehicle collision or single-vehicle accident. In addition, head-on/sideswipe opposite direction collisions and single-vehicle rollover crashes were identified. If an accident involved a sequence of two or more events (as could be ascertained in the Utah data), collision with another vehicle took precedence over a single vehicle event, head-on/sideswipe opposite direction collision took precedence over other types of collisions, and rollover took precedence over other single-vehicle events. Counts of each of these types of crashes were made for each roadway section for use in calculating the rates.

Statistical Methods

The accident rate per 100 million vehicle miles traveled for an individual road section was calculated as

$$R = (Y/VM) * 10^8$$

where

R = observed rate,
Y = observed number of accidents,
VM = vehicle miles of travel calculated as $ADT * 365 * T * L$,
ADT = average daily traffic (vehicles per day),
T = number of years over which accidents were counted, and
L = section length (mi) (1 mi = 1.61 km).

Accident rates corresponding to all accidents, serious injury accidents (A or K), injury accidents (C, B, A, or K), PDO accidents, multivehicle accidents, head-on or sideswipe opposite direction accidents, single-vehicle accidents, and single-vehicle rollover accidents have been analyzed using regression models. The specific aims of the modeling process were to obtain standard errors and confidence intervals for estimated accident rates and to determine whether the observed reduc-

tion in the crude accident rates for wider medians persisted after adjusting for other roadway variables.

A log-linear regression model was used to simultaneously assess the effects of median width and several other roadway variables on the accident rate. This model may be represented algebraically as

$$\log(\lambda) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

where

λ = expected value of $R = E(R) = [E(Y)/VM] * 10^8$ (log denotes logarithm to base e), and

X_i = indicator (dummy) variables for categorical roadway characteristics (e.g., functional class) or actual values for quantitative roadway characteristics (e.g., right shoulder width).

Note that $\exp(\beta_i)$ (i.e., e^{β_i}) represents the relative effect of a unit change in X_i on the accident rate.

Log-linear models assume that the effect of variables on the accident rate is multiplicative rather than additive as in linear models. Estimated rates from log-linear models cannot be negative. Log-linear models have been widely used in statistical analyses of count data [see McCullagh and Nelder (*9*) and references therein] and have recently been used in transportation studies by Joshua and Garber (*10*) for truck accident rates, Hauer and Persaud (*11*) for railway-crossing accident rates, and Zegeer et al. (*12*) for highway accident rates. Zegeer et al. (*12*) considered both additive and multiplicative (i.e., log-linear) models and concluded that the multiplicative models provided a better fit to the data.

To obtain estimates, standard errors, and confidence intervals, the negative-binomial variance function was assumed for the accident count per section, that is,

$$Var(Y) = E(Y) + K * [E(Y)]^2$$

where K has the same value for all sections and $Var(Y)$ and $E(Y)$ are the variance and expected value, respectively. The classical distribution for accident counts is the Poisson distribution for which the variance is equal to the mean. However, variances in excess of the mean are often observed (*13*), partly because not all relevant variables are included in the model. The negative binomial distribution is a natural extension of the Poisson, which accounts for this excess variability and has certain desirable theoretical properties (*14*). The negative-binomial distribution for accident counts has been used recently by Hauer and Persaud (*11*) and Hauer et al. (*15*), and these authors have validated its use for transportation studies. Maher (*16*) also used the negative binomial distribution to explain traffic accident migration and states that "it has become standard" to use this distribution. This assumption was validated in our study by calculating the mean and variance of Y (for total accidents) for homogeneous subgroups of road sections and plotting the variance against the mean.

The beta coefficients in the regression model were estimated by the method of quasi-likelihood, and the value of K was estimated by the method of moments (*9*). Others have used maximum likelihood estimation (*11,15*), but it has been suggested that quasi-likelihood estimation for the beta coefficients and the method of moments for K is a more robust estimation procedure (*14*) and therefore have been used here. The estimation procedures were carried out using the statistical package GLIM (*17*), and the GLIM macros (or procedures) for fitting these models are given by Breslow (*13*). For Utah, the estimated value of K was about 0.6 and for Illinois it was about 1.4, suggesting that accident rates for similar sections of highway are more variable in Illinois. This is most likely due to greater variability in driver and environmental conditions in Illinois than in Utah.

In the regression models, median width has been examined both as a categorical variable (six categories for Utah and eight categories for Illinois) and as a continuous variable in the form of a quartic (fourth-degree) polynomial function without a linear term, because this particular function closely resembled the observed rates. When median width has a categorical representation, no trend is assumed, whereas the continuous representation adopted in this study assumes a quartic polynomial trend on the log scale for the accident rates. As in all continuous forms of modeling, the data are "smoothed" by the assumed trend. By using both representations, comparison of the estimated rates (and confidence intervals) for the categories allows a check on the appropriateness of the form of the assumed trend in the continuous model. In all cases the trends were consistent with a quartic polynomial trend. For comparison purposes, in this paper results for both forms of representation are reported.

The purpose of the analysis was to determine the effect of median width on the accident rate after controlling or adjusting for other variables. Variables that have been controlled by design through restricting the analysis to particular (homogeneous) sections were listed earlier. Variables included in the regression models are functional classification (rural–Interstate/freeway, rural other, urban–Interstate/freeway, urban other), posted speed limit, right shoulder width (continuous), access control (none, partial, full—Illinois only), curvature (dichotomous as described above—Utah only), log (average daily traffic) (continuous), and log (section length) (continuous). Section length was included as a surrogate for other variables not included that may be correlated with section length. Because the sections were constructed to be homogeneous, shorter sections occur where the roadway characteristics are changing more rapidly.

Many of the variables included in the regression model were correlated with median width, and several combinations of median width and other variables had very few or no sections. For example, Interstate road sections had larger median widths, whereas other functional classes had smaller median widths, although there was some overlap. This made the fitting of interactions between median width and other variables difficult. Where possible, such interactions were examined, but no significant interactions were found.

The estimated effects of median width obtained from these models (especially those with a categorical representation) may be conservative, since when variables correlated with median width are included in the models, they will absorb some of the effect of median width. For example, if functional class is omitted from the model, the effect of median width increases and vice versa. Inclusion of such variables has been done deliberately so that any median width effects detected cannot be attributed to other confounding variables.

RESULTS

Table 1 gives the characteristics of the road sections that have been used in the accident rate analyses. Because there were fewer sections in the Utah data, only six median width categories were used rather than eight as for Illinois. Note also that there were very few sections in the Utah data with median width in the range 30 to 54 ft (9.2 to 16.5 m) and very few sections with functional classification as urban–Interstate/freeway.

The crude average accident rates by median width for total accidents and severity and collision types are given in Table 2. The total accident rate appears to decline steadily with increasing median width. For Utah it declines from 650 for sections with no median to 111 accidents per 100 million vehicle-mi (179 accidents per 100 million vehicle-km) traveled for sections with median width at least 85 ft (25.9 m). Thus the crude total accident rate is reduced by a factor of about 6 over this range of median width. The decrease in the total accident rate for Illinois declines by a factor of about 13.

Serious injury (i.e., AK), all injury (CBAK), and property-damage-only accidents also show many-fold reductions over this range of median width. The rate for multivehicle accidents declines steadily with increasing median width, and head-on/sideswipe opposite direction accidents in particular show a dramatic decrease with increasing median width. On the contrary, the rates for single-vehicle accidents (Utah) and single-vehicle rollover accidents in particular show little relationship to median width.

The many-fold reductions observed in these accident rates cannot all be attributed to the effect of median width because of confounding by other variables. It is for this reason that the models including these confounding factors are developed. The relative effect of median width on the total accident

TABLE 1 Number of Sections (N), Number of Roadway Miles (Miles) with Various Characteristics for Utah and Illinois

Utah			Illinois		
Category	N	Miles	Category	N	Miles
Overall	982	973.8		2481	2081.3
Median Width (ft)					
0	176	68.7	0	567	219.0
1-10	257	110.9	1-24	199	67.0
11-29	213	114.7	25-34	176	89.4
30-54	52	76.8	35-44	479	304.2
55-84	179	298.7	45-54	200	139.7
85-110	105	303.9	55-64	450	538.4
			65-84	239	424.6
			85-110	171	298.9
Functional Class					
rur_int	284	653.0	rur_int	846	1293.8
rur_oth	130	73.5	rur-oth	343	182.0
urb_int	64	43.9	urb_int	436	279.9
urb_oth	504	203.3	urb_oth	856	325.6
Speed Limit					
35-40	183	61.6	35-40	370	128.8
45-50	118	44.7	45-50	348	124.1
55	146	101.8	55	889	486.0
65	305	663.9	65	874	1342.4
missing	230	101.7			
Right Shoulder Width (ft)					
0	315	119.2	0	401	155.0
1-5	121	62.3	1-5	65	25.6
6-10	495	768.5	6-10	1406	1223.0
11-23	51	23.9	11-23	609	677.6
Curvature > 1 Degree					
no	756	605.6		NA	
yes	226	368.2			
Access Control					
none	NA		NA	872	356.8
partial				435	216.9
full				1174	1507.6

NOTE: 1 mi. = 1.61 km, 1 ft. = 0.305 m

TABLE 2 Crude Average Accident Rates per 100 Million Vehicle-mi and Estimated Relative Effects of Median Width on the Total Accident Rate [Median Width Is Represented Both as a Categorical Variable and as a Continuous Variable, Adjusting for Functional Class, Posted Speed Limit, Right Shoulder Width, Access Control (Illinois Only), Curvature (Utah Only), Log (ADT) and Log (Section Length)]

Median Width (MW)			Average Accident Rate (R)								Relative Effect on Total Accident Rate		
											Categorical		Continuous
Category	Mean	N	AK	CBAK	PDO	MVeh	SVeh	HO	Roll	Total	Estimate	95% Conf Int	Estimate
Utah													
0	0.0	176	48	220	430	522	127	10	14	650	1.00	(1.00, 1.00)	1.00
1-10	9.4	257	47	203	416	521	97	10	5	618	1.06	(0.89, 1.26)	0.96
11-29	14.9	213	45	159	303	373	89	8	7	462	0.97	(0.81, 1.17)	0.91
30-54	46.3	52	19	53	106	51	109	1	29	159	0.56	(0.36, 0.88)	0.61
55-84	71.7	179	20	42	95	31	106	1	22	137	0.51	(0.34, 0.78)	0.52
85-110	101.0	105	22	44	67	18	93	0	29	111	0.47	(0.30, 0.71)	0.47
ALL	32.0	982	38	142	282	321	103	6	14	424			
Illinois													
0	0.0	567	46	214	477	605	86	21	5	692	1.00	(1.00, 1.00)	1.00
1-24	12.8	199	40	194	452	578	69	12	8	647	1.05	(0.86, 1.28)	0.96
25-34	29.8	176	42	115	177	200	92	3	15	292	0.81	(0.62, 1.06)	0.84
35-44	39.7	479	16	48	82	78	51	2	6	129	0.77	(0.59, 0.99)	0.76
45-54	49.2	200	10	37	90	66	61	2	7	127	1.00	(0.74, 1.35)	0.69
55-64	63.8	450	5	14	31	18	27	1	3	45	0.62	(0.46, 0.83)	0.62
65-84	71.9	239	7	18	41	19	40	1	5	59	0.64	(0.46, 0.88)	0.60
85-110	88.9	171	6	20	34	18	36	1	6	53	0.68	(0.48, 0.96)	0.65
ALL	39.4	2481	22	91	193	226	58	7	6	283			

NOTE: 1 ft. = 0.305 m

Mean = average median width
N = number of road sections
Total = overall accident rate
AK = A+K rate
CBAK = all injury rate
PDO = property damage only
MVeh = multi-vehicle accident rate
SVeh = single-vehicle accident rate
HO = head-on/sideswipe opposite direct. rate
Roll = single-vehicle rollover rate

rate after adjustment for other variables via the log-linear regression model is also given in Table 2 and shown graphically in Figure 1. The estimate and standard error of the coefficients for fitted log-linear models showing the continuous effect of median width and the other independent variables are presented in Table 3.

The continuous estimates given in Table 2 were obtained by inserting the average median width for each category into these equations. The interpretation of these relative effects is that, when all the other variables are the same and the only difference is the median width, the relative effect describes the proportional reduction in the total accident rate. For example, using the Illinois equation (continuous), the total accident rate for an average median width of 40 ft (12.2 m) is about 76 percent of the rate for median width zero (no median), and for an average median width of 64 ft (19.5 m) (see Table 3 for mean of interval) it is 62 percent. An estimate of the safety benefit of increasing the median from 40 to 64 ft (12.2 to 19.5 m) is obtained as $(0.62 - 0.76)/0.76 = -0.18$. Therefore, one would expect an 18 percent reduction in the

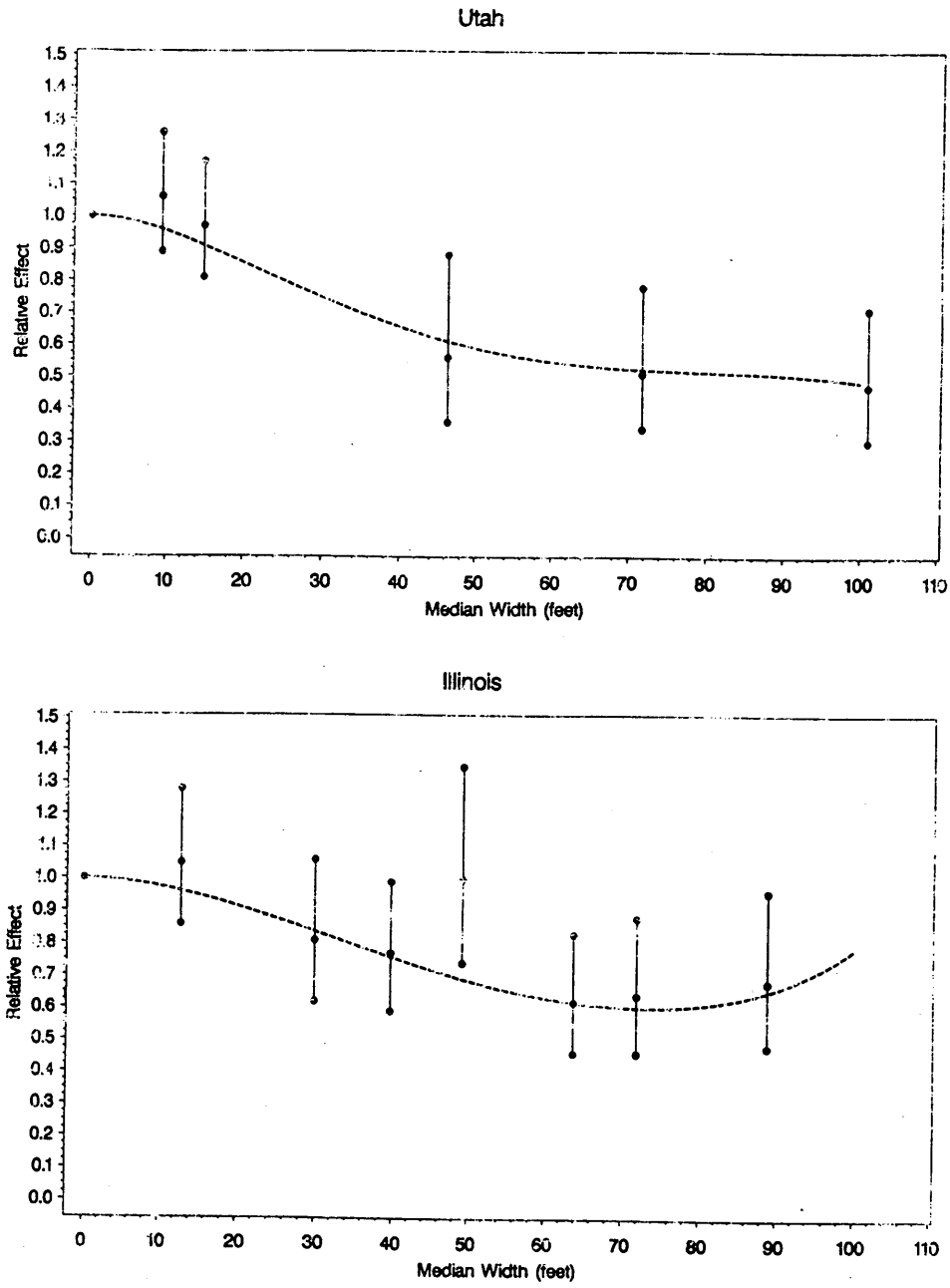


FIGURE 1 Estimated relative effects of median width on the total accident rate when median width is represented both as a categorical variable and as a continuous variable, adjusting for functional class, posted speed limit, right shoulder width, access control (Illinois only), curvature (Utah only), log (ADT) and log (section length). Note: 1 ft = 0.305 m.

TABLE 3 Fitted Log-Linear Regression Models for Total Accident Rate Showing Continuous Effect of Median Width and Other Variables

	Parameter	Estimate	Standard Error
UTAH			
	Constant	6.196	0.2943
	Median width^2	-5.589×10^{-4}	3.549×10^{-4}
	Median width^3	8.940×10^{-6}	7.083×10^{-6}
	Median width^4	-4.105×10^{-8}	3.716×10^{-8}
	Rural other vs rural interstate	-1.078	0.2757
	Urban interstate vs rural interstate	-0.2911	0.1714
	Urban other vs rural interstate	-0.5081	0.2782
	Curvature > 1 degree	0.0456	0.0754
	Right shoulder width	-0.0352	0.0082
	Speed limit 45-50 vs 35-40	0.5187	0.1097
	Speed limit 55 vs 35-40	0.4679	0.1149
	Speed limit 65 vs 35-40	-0.5417	0.2015
	Speed limit missing vs 35-40	0.6461	0.1041
	Log (average daily traffic)	-0.1389	0.0448
	Log (section length)	-0.1962	0.0308
Illinois			
	Parameter	Estimate	Standard Error
	Constant	4.587	0.1655
	Median width^2	-2.622×10^{-4}	2.397×10^{-4}
	Median width^3	2.062×10^{-6}	5.799×10^{-6}
	Median width^4	3.167×10^{-9}	3.740×10^{-8}
	Rural other vs rural interstate	0.4293	0.1308
	Urban interstate vs rural interstate	-0.0566	0.0975
	Urban other vs rural interstate	0.7921	0.1368
	Access control partial vs none	0.3723	0.1298
	Access control full vs none	0.4546	0.1280
	Right shoulder width	-0.0460	0.0110
	Speed limit 45-50 vs 35-40	0.5541	0.1140
	Speed limit 55 vs 35-40	0.5121	0.0962
	Speed limit 65 vs 35-40	-0.5434	0.1000
	Log (average daily traffic)	-0.2509	0.0495
	Log (section length)	-0.1232	0.0251

accident rate. On the other hand, if one reduced an existing median of 64 ft (19.5 m) to a median of 40 ft (12.2 m), one would expect a 23 percent increase in the total accident rate [(0.76 − 0.62)/0.62 = 0.23].

Thus the decline in the crude total accident rates with increasing median width given in Table 2 persists, albeit it to a lesser degree, after adjustment for these other confounding variables. Similar trends are shown for Utah and Illinois. These results indicate that there is little reduction in the accident rate for median widths up to about 25 ft (7.6 m). Whereas this lack of decrease is not as apparent in the smoothed continuous models, the categorical estimates for the smaller median widths are a little greater than 1.0 (indicating no difference from a median width of zero). The decline in accident rate, particularly in the categorical model, is most apparent for median widths beyond about 20 to 30 ft (6.1 to 9.2 m). The decreasing trend seems to become level at median widths of approximately 60 to 80 ft (18.3 to 24.4 m), particularly for Illinois.

The estimated relative effects for serious injury, all injury, and property-damage-only accident rates are given in Table 4. Logic suggests that the effect should be stronger for more severe accidents because wider medians would reduce the likelihood of collisions between vehicles traveling in opposite directions, which tend to have serious injury consequences. However, although the effect of median width on the accident rate is slightly stronger for injury accidents (but not AK accidents) than for property-damage-only accidents for Utah, the effect appears to be much the same for all severity classes for Illinois.

The estimated relative effects (continuous model) for multivehicle, single-vehicle, head-on/sideswipe opposite direction, and single-vehicle rollover accident rates are shown in Figure 2. For Utah the effect of median width is very similar for multivehicle and single-vehicle accidents, whereas for Illinois the effect is larger for multivehicle accidents, as might be expected intuitively.

More specifically, one might expect that median width would have its most dramatic effect on head-on/sideswipe opposite direction accidents. This is demonstrated clearly by the Illinois data. However, for Utah, although median width appears to have a dramatic effect on head-on/sideswipe opposite direction accidents after about 40 ft (12.2 m), the size of the effect is somewhat similar to the effect for multivehicle accidents in general.

Median width had little effect on single-vehicle rollover accidents for Illinois but appeared to have a rather sizable effect for Utah.

TABLE 4 Estimated Relative Effects of Median Width on Serious Accident Rates (AK), Injury Accident Rates (CBAK), and Property-Damage-Only Accident Rates (PDO) [Uses Models in Which Median Is Represented Both as a Categorical (cat) and as a Continuous (cts) Variable, Adjusting for Functional Class, Posted Speed Limit, Right Shoulder Width, Access Control (Illinois Only), Curvature (Utah Only), Log (ADT), and Log (Section Length)]

	Median Width (Mean)	AK cat	AK cts	CBAK cat	CBAK cts	PDO cat	PDO cts
UTAH							
	0 (0.0)	1.00	1.00	1.00	1.00	1.00	1.00
	1-10 (9.4)	0.95	0.96	0.92	0.94	1.10	0.97
	11-29 (14.9)	1.01	0.91	0.91	0.86	0.99	0.92
	30-54 (46.3)	0.65	0.63	0.53	0.50	0.56	0.65
	55-84 (71.7)	0.62	0.57	0.48	0.46	0.51	0.52
	85-110 (101.0)	0.73	0.66	0.57	0.52	0.41	0.42
ILLINOIS							
	0 (0.0)	1.00	1.00	1.00	1.00	1.00	1.00
	1-24 (12.8)	1.00	0.97	1.04	0.98	1.07	0.95
	25-34 (29.8)	1.10	0.86	0.95	0.89	0.76	0.81
	35-44 (39.7)	0.88	0.77	0.84	0.80	0.72	0.74
	45-54 (49.2)	0.84	0.68	0.97	0.72	1.04	0.69
	55-64 (63.8)	0.60	0.57	0.60	0.61	0.63	0.64
	65-84 (71.9)	0.68	0.54	0.67	0.58	0.63	0.64
	85-110 (88.9)	0.58	0.57	0.64	0.64	0.70	0.67

NOTE: 1 ft. = 0.305 m

The results for head-on/sideswipe opposite direction and for rollover accidents should be interpreted with some caution, especially for Utah, because there were very few accidents of these types. For Illinois, 1,980 sections (out of a total of 2,481) and, for Utah, 699 sections (out of a total of 982) had no head-on/sideswipe opposite direction accidents, whereas 2,241 sections in Illinois and 907 sections in Utah had no single-vehicle rollover accidents.

CONCLUSIONS

This investigation represents an attempt to define the relationship between median width and accident rate while controlling for other confounding variables. Although there were some studies in the prior literature relating to median width, in general the literature on this subject is quite sparse. Thus, there is little available information on an issue that is even more critical today given the current movement toward adding lanes to multilane facilities to enhance capacity without purchasing additional right-of-way. Thus, even with the caveats stated below, this study is a beginning point in the development of much needed information related to median width and safety.

This study has the advantage of a more comprehensive data base than prior studies. In addition, the data used here are more current than the data in the older studies, and we were able to use data from two states rather than only one, which allowed us to look at consistency of findings between the states. Furthermore, there is greater mileage of four-lane divided highway and thus miles of median in each of the study states than had been the case in earlier studies, along with a wider range of median widths.

There are, however, some necessary caveats that must be stated. First, in any study that attempts to control for confounding variables through statistical means rather than through the design of the study (i.e., by actually assigning different median widths to similar sections of the highways), the validity of the results depends on how well the confounding variables are identified and measured. Whereas we attempted to control for major confounding variables in the analyses conducted here, there are clearly other variables that were either not measured in our data base or not used in the final model simply because of the need to limit the model to as few variables as possible. These possible confounding variables include vertical grade, median slope, type of traffic (e.g., percent heavy trucks), environmental factors, additional geometric variables related to details of curvature or sideslope design, and general exposure factors. Even with these caveats, the results are important.

The general findings indicate that accident rates decrease with increasing median width, even when other confounding variables are controlled for. Whereas the degree of improvement due to median width was not exactly the same in the Utah data as in the Illinois data, the same general trends were observed in the two states. Second, it was also apparent that there was very little decrease, if any, in the various accident rates for medians less than approximately 20 to 30 ft (6.1 to 9.2 m) in width in the two states. Thus, in terms of modification of existing roadways, this finding indicates that decreasing any median width that is greater than 20 to 30 ft (6.1 to 9.2 m) to 30 ft (9.2 m) or less to enhance capacity would probably be accompanied by a decrease in the level of safety on the roadway. [Unfortunately, we could not determine the exact "breakpoint" where the safety effect ends. Whereas the categorical data from both states indicated no safety effect

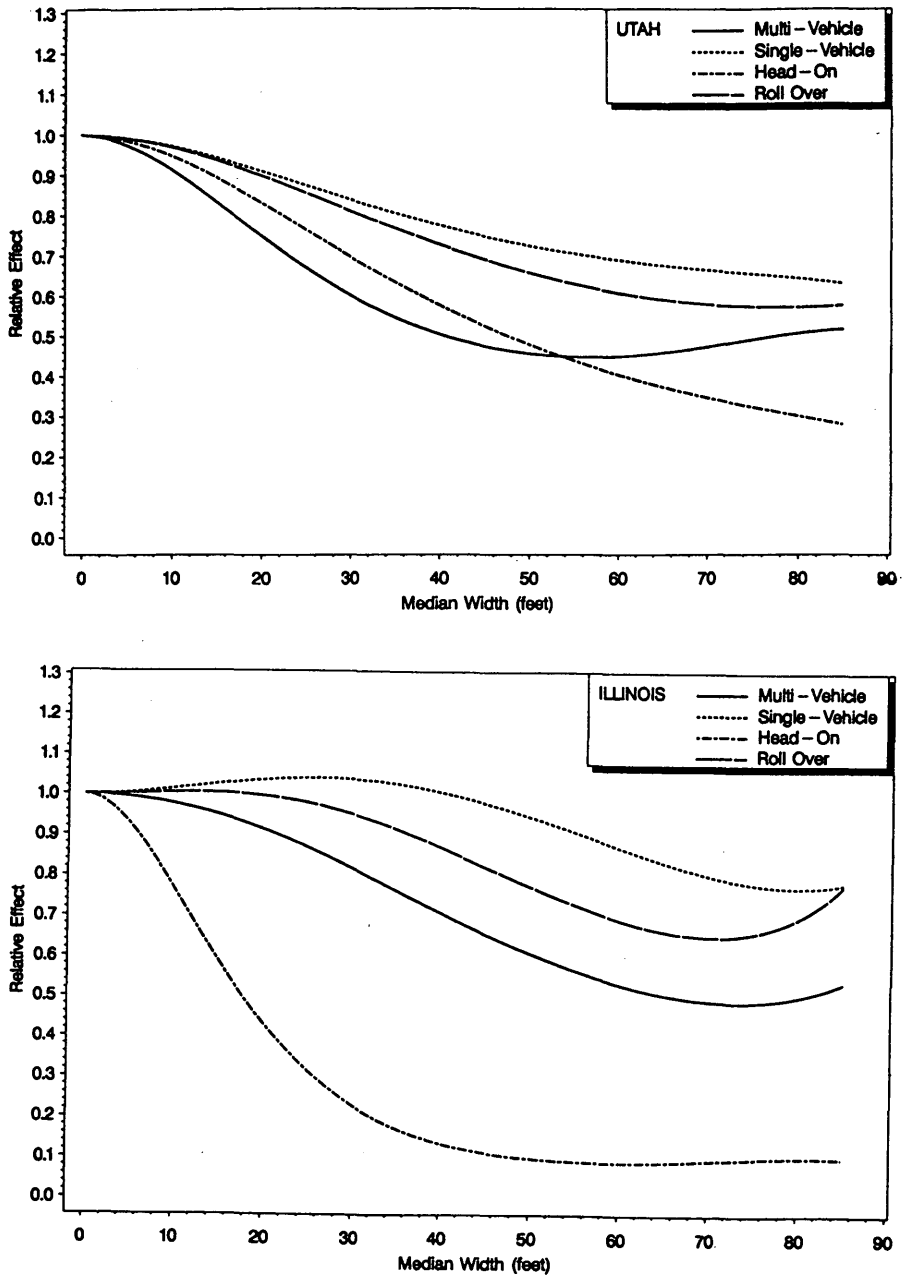


FIGURE 2 Estimated relative effects of median width on multivehicle accident rates, single-vehicle accident rates, head-on/sideswipe opposite direction accident rates, and single-vehicle rollover accident rates from models in which median width is represented as a continuous variable, adjusting for functional class, posted speed limit, right shoulder width, access control (Illinois only), curvature (Utah only), log (ADT), and log (section length). Note: 1 ft = 0.305 m.

for medians less than approximately 20 to 25 ft (6.1 to 7.6 m), there were not adequate numbers of 20-ft (6.1-m), 25-ft (7.6-m), or 30-ft (9.2-m) medians to allow separate analyses of these individual categories.]

There were also differences noted from what might have been traditionally hypothesized as the manner in which median width affects safety. For example, it might have been hypothesized that median width would be primarily related to decreases in "crossover accidents" involving head-on crashes between opposing vehicles. As a result of reducing these crossover accidents, changes in median width might have been expected to have a much greater effect on severe crashes than on less severe or property-damage-only crashes. We did not find either to be the case. As noted above, whereas we found significant changes in head-on crashes in both states, the changes in head-on crashes were only a small part of the overall decrease in total multivehicle accidents in each state. In addition, we did not find much difference in the effects of width on

accident severity—the less severe crashes were affected as much as the more severe.

However, these results are not as surprising as first thought if viewed under the earlier-stated modified assumption of how medians affect safety. If instead of just acting as a buffer between vehicles that run off the road left toward each other, it is assumed that a median may well be serving as an escape area or clearzone for vehicles that are avoiding possible crashes with vehicles in their own lanes, one would see decreases in multivehicle crashes of all types (even rear-ends) and perhaps increases (or no change) in single-vehicle accidents due to the additional "roadside" to run off into. This is indeed what we found in the data—clear decreases in multivehicle crashes of all types and lesser or no decreases in the single-vehicle ran-off-road type crash.

Thus, in summary, it may be that we need to view the median differently, and this new view may affect median design. If the median is to "sell itself" to the driver as a safe escape area, it must clearly be wide enough to give the motorist the perception of safety. If the median is so narrow that heavy oncoming traffic on the opposing roadway reduces the perception of additional safety, it will not be used as much, and accident reductions will decrease.

A major point of interest is how these findings agree with design guidelines provided in the AASHTO *Policy on Geometric Design* (*1*). It is difficult to summarize AASHTO median-width and barrier-need guidelines, since material is found in a variety of sections of the *Policy* and because "hard" guidelines are not presented. This is due, of course, to the lack of hard data on the issue.

The general guideline provided is that careful study is needed of all locations. With respect to rural arterials, it appears that the policy suggests that medians of 60 ft (18.3 m) or more should be provided whenever feasible. In locations with restricted right-of-way, medians of 30 ft (9.2 m) or more are recommended. However, the additional information related to median width at intersections on rural arterials confuses the issue somewhat. Here, it is suggested that median widths of 12 to 30 ft (3.7 to 9.2 m) function quite well in that they provide room for turn lanes and, thus, protect turning vehicles; that median widths of 30 to 50 ft (9.2 to 15.3 m) may be suitable if detailed study of operational characteristics of the traffic are conducted; but that medians of 50 to 80 ft (15.3 to 24.4 m) ". . . have developed accident problems in some cases. . ." Thus, the designer is left with the impression that wider medians should not be used in places where at-grade intersections are present.

With respect to urban freeways, the general guideline is again to use medians that are as wide as possible. On four-lane facilities in areas of restricted right-of-way, it is suggested that 10-ft (3.1-m) medians are acceptable as long as a positive barrier is used. For six-lane facilities, a minimum width of 22 to 26 ft (6.7 to 7.9 m) is acceptable, again as long as a barrier is used. It is also interesting to note that a 50-ft (15.3-m) median is shown as a typical (nonbarrier) median width in a figure depicting a typical cross section with a median.

With respect to rural freeways, even less guidance is given. It is noted that 50- to 90-ft (15.3- to 27.5-m) medians are common. In sketches of typical cross section, a 50-ft (15.3-m) median is shown. It is further noted that in suburban areas, restricted right-of-way may lead to medians in the range of 10 to 30 ft (3.1 to 9.2 m) and that in these cases "median barrier is usually warranted as a safety measure."

Given the "softness" of the guidelines presented in the AASHTO policy, it is difficult to say whether the findings of this study support the design policy presented there. In this study, we find evidence that medians that are 50 ft (15.3 m) wide are indeed much safer than the no-median or narrow median condition. However, we also find that even wider medians [up to 80 ft (24.4 m) or more] appear to provide even greater safety benefits. If one takes literally the advice provided by the AASHTO guidebook concerning the need for barriers on either 10-ft (3.1-m) or 20- to 26-ft (6.1- to 7.9-m) medians, one might assume that four-lane medians greater than 15 ft (4.6 m) in width might be acceptable without barriers. Our findings do not support this at all. Indeed, the data here indicate that one needs to have a median at least 20 to 30 ft (6.1 to 9.2 m) in width before any safety effect is seen and that there are significant increases in the level of safety as one moves from 30 ft (9.2 m) to the wider median widths. Thus, in the design of new highways, our findings would support medians considerably wider than 30 to 40 ft (9.2 to 12.2 m).

This same information can be used in a slightly different way to provide information to the designer who is looking at the situation of potential lanes being added within the median. The conclusion from these data would be that safety benefits will indeed be lost by narrowing a median to any extent, and that if the median is narrowed to a width of between 20 ft (6.1 m) and 30 ft (9.2 m) (or less), essentially all of the safety benefit of the median may be lost unless a positive barrier is used. Unfortunately, because of the lack of barrier sections in the data set, we could not analyze the question of the benefit of placing positive barriers in the median.

In terms of needed additional research, it appears that these data have provided new information with respect to width of nonbarrier medians and the effects on safety—medians wider than approximately 25 to 30 ft (7.6 to 9.2 m) have a significant safety benefit, and the wider the median the better, up to approximately 65 to 80 ft (19.8 to 24.4 m). However, the most obvious remaining gap in knowledge is when to install positive barriers. At what width do the benefits of reductions in severe (cross-median) crashes outweigh the increase in less severe crashes? To conduct such a study will require a large sample of medians of various widths [at least in the range of 0 to 50 ft (0 to 15.3 m)] with and without barriers—clearly a multistate study.

REFERENCES

1. *A Policy on Geometric Design of Highways and Streets*. American Association of State Highway and Transportation Officials, Washington, D.C., 1984.
2. F. W. Hurd. Accident Experience with Traversable Medians of Different Widths. *Bulletin 137*, Highway Research Board, National Research Council, Washington, D.C., 1957, pp. 18–26.
3. E. T. Telford and R. J. Israel. Median Study (California). *Proc.*, 32nd Annual Meeting, Highway Research Board, 1953, pp. 208–231.
4. J. R. Crosby. Cross-Median Accident Experience on the New Jersey Turnpike. *Bulletin 266*. Highway Research Board, National Research Council, Washington, D.C., 1960, pp. 63–77.
5. C. E. Billion and N. C. Parsons. Median Accident Study—Long Island, New York. *Bulletin 308*, Highway Research Board, National Research Council, Washington, D.C., 1962, pp. 64–79.

6. G. R. Garner and R. C. Deen. Elements of Median Design in Relation to Accident Occurrence. In *Highway Research Record 432*, HRB, National Research Council, Washington, D.C., 1973, pp. 1–11.
7. F. M. Council and E. G. Hamilton. *Highway Safety Information System Guidebook for the Utah State Data Files.* Highway Safety Research Center, Chapel Hill, N.C., 1990.
8. F. M. Council and C. Williams. *Highway Safety Information System Guidebook for the Illinois State Data Files.* Highway Safety Research Center, Chapel Hill, N.C., 1990.
9. P. McCullagh and J. A. Nelder. *Generalized Linear Models* (2nd edition). Chapman and Hall, New York, 1989.
10. S. C. Joshua and N. J. Garber. Estimating Truck Accident Rate and Involvements Using Linear and Poisson Regression Models. *Transportation Planning and Technology*, Vol. 15, 1990, pp. 41–58.
11. E. Hauer and B. N. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1987, pp. 131–140.
12. C. V. Zegeer, D. W. Reinfurt, J. Hummer, L. Herf, and W. Hunter. Safety Effects of Cross-Section Design for Two-Lane Roads. In *Transportation Research Record 1195*, TRB, National Research Council, Washington, D.C., 1988, pp. 20–32.
13. N. E. Breslow. Extra-Poisson Variation in Log-Linear Models. *Applied Statistics*, Vol. 33, 1984, pp. 38–44.
14. J. F. Lawless. Negative Binomial and Mixed Poisson Regression. *The Canadian Journal of Statistics*, Vol. 15, 1987, pp. 209–225.
15. E. Hauer, B. N. Persaud, A. Smiley, and D. Duncan. Estimating the Accident Potential of an Ontario Driver. *Accident Analysis and Prevention*, Vol. 23, 1991, pp. 133–152.
16. M. J. Maher. A Bivariate Negative Binomial Model to Explain Traffic Accident Migration. *Accident Analysis and Prevention*, Vol. 22, 1990, pp. 487–498.
17. R. J. Baker and J. A. Nelder. *The Generalized Linear Interactive Modelling System, Release 3.77.* Numerical Algorithms Group, London, 1987.

DISCUSSION

SHAW-PIN MIAOU
Center for Transportation Analysis, Energy Division, Oak Ridge National Laboratory, P.O. Box 2008, MS 6366, Building 550DA, Oak Ridge, Tenn. 37831.

The authors examined the effect of median width on vehicle accident rate for multilane divided highway sections with a traversable or nonbarrier median. Log-linear regression models with a negative-binomial variance function were used to study the effect. The authors should be commended for addressing a very important, yet difficult, problem. Overall, this is a well-written paper that presents some interesting empirical results. However, some of the results seem to be questionable:

1. This study failed to separate paved inside shoulders from the rest of the median. Paved inside shoulders are part of the roadway immediately contiguous with the traveled way and are important features of divided multilane highways. Failing to consider "paved inside shoulder width" in this study posted two potential problems: (*a*) the model results on the effect of median width are difficult to interpret in a design context and (*b*) it is entirely possible that the paved inside shoulder width was associated with the accident rate, not the rest of median width. To illustrate, let the paved inside shoulder width be X_1 and the rest of median width be X_2. In addition, let the total median width be X ($= X_1 + X_2$) and the number of accidents be Y. Furthermore, assume that X_1 is correlated with Y and X_2 is independent of Y. We can show that the correlation coefficient of Y and X, denoted by ρ_{xy}, does not vanish and can be computed as $\rho_{xy} = \mathrm{Cov}(X_1, Y)/[\mathrm{Var}(Y)\mathrm{Var}(X)]^{1/2}$.

2. (Table 1) Does "right shoulder width" include both the width of paved and unpaved shoulders? It does not seem reasonable to have road sections with a right shoulder width of 23 ft. Two related questions are as follows: How many road sections have a right shoulder width of 13 ft or more? Were these road sections particularly influential in estimating model coefficients?

3. (Table 2) Many rural Interstate road sections in the Utah roadlog file were coded as having a median width of 99 ft, which really meant that the road section's median width was equal to or greater than 99 ft. How did the authors handle these road sections?

4. (Table 4) The estimated coefficients for ADT having an algebraic sign contrary to expectation. The estimated coefficient for log(ADT) was −0.1389 in the Utah model and −0.2509 in the Illinois model. Thus, both models indicated that, for road sections of a particular functional class (and speed limit and access control), as ADT increased, total accident rate decreased. This result is apparently not acceptable. One possible reason for this to occur is that ADT alone did not give a good description of the traffic condition. Variables related to highway capacity, such as the number of lanes, should be considered in the model. Another possible reason is the collinearity problem to be discussed later.

5. (Table 4) The estimated regression coefficients for "median width" have very low *t*-statistics, indicating that the effect of median width on accident rate was poorly determined from the data. For the Utah model, the *t*-statistics of the estimated coefficients for $(\text{median})^3$ and $(\text{median})^4$ were about 1.26 and −1.10, respectively. For the Illinois model, *t*-statistics of the estimated coefficients for $(\text{median})^2$, $(\text{median})^3$, and $(\text{median})^4$ were about −1.09, 0.36, and 0.08, respectively. These low *t*-statistics were indications to the authors that they might have "oversmoothed" or "overinterpreted" the data. Therefore, the statements in this paper on the effect of median width, such as that the decreasing trend seems to become level at median widths of approximately 60 to 80 ft, particularly for Illinois, are questionable. Why not just consider the first- and the second-order terms [i.e., (median) and $(\text{median})^2$]?

6. (Table 4) Some of the variables considered in the model were extremely collinear (e.g., functional class, speed limit, and access control were highly correlated with one another). This collinearity problem may have made the interpretation of the fitted log-linear regression models difficult and the results questionable.

Some examples of Item 6 are as follows:

- Unreasonable speed limit effect?—If we use the fitted models to shed some light on the effect of speed limit change (from 55 to 65 mph) in 1987 on accident rates for rural Interstate highway sections, we would find that the models suggested a 64 and a 65 percent reduction in total accident rate for Utah and Illinois, respectively. These results cannot be supported by any highway statistics. This is probably a result of the distortion produced by the collinearity of some of the

covariates. The computation of these reductions can be carried out as follows: Take Utah for example. Let the accident rates of any rural Interstate road section before and after the speed limit change be λ_{55} and λ_{65}, respectively. Provided that everything else was the same, the fitted model suggested that the ratio of these two accident rates would be $\lambda_{65}/\lambda_{55}$ = exp(−0.5417)/exp(0.4679) = exp(−1.0096) = 36 percent. Therefore, according to the model, the drop in accident rate on a rural Interstate section as a result of the speed limit change would have been 64 percent.

• Unexpected signs in coefficients for functional class variables?—For Utah, the estimated coefficients for functional class variables (i.e., "rural other versus rural Interstate," "urban Interstate versus rural Interstate," and "urban other versus rural Interstate") were negative (i.e., −1.078, −0.2911, and −0.5081, respectively). The negative sign also appeared in the Illinois model for "urban other versus rural Interstate." If we disregard other variables and focus on functional class variables alone, the Utah model suggests that rural other highways, urban Interstates, and urban other highways had a lower total accident rate than that of rural Interstates, which was contrary to what one would usually expect. But because functional class, speed limit, and ADT are highly correlated with one another, it may not be appropriate to examine functional class variables alone. The authors should make this clear in the paper.

Now, consider two hypothetical road sections in Utah: one rural and one urban Interstate section. Assume that these two road sections have the same geometric design characteristics, section length, and speed limit. Furthermore, assume that the rural and urban road sections have an ADT of 5,000 and 50,000 vehicles, respectively. Then, according to the model, the ratio of the accident rate between these two road sections is $\lambda_{urban}/\lambda_{rural} = \exp\{-0.2911 - 0.1389 \times [\log_e(50{,}000) - \log_e(5{,}000)]\} = \exp(-0.611) = 54$ percent. That is, the total accident rate of the urban road section is 46 percent lower than that of the rural road section. It is arguable that this ratio does not seem to be reasonable. More information will be needed for the readers to make a better judgment on this. For example, the authors may want to (*a*) cross-classify the number of road sections by functional class, speed limit, access control, and ADT in Table 1 and (*b*) tabulate the accident rate by functional class, speed limit, access control, and ADT.

• Unexpected signs in coefficients for "access control" variables?—For Illinois, the estimated coefficients for "access control partial versus none" and "access control full versus none" were 0.3723 and 0.4546, respectively. This implies that for any road section, the tighter the access control we apply to it, everything else being the same, the higher the accident rate would be, which is unreasonable. Again, to make a better judgment on the reasonableness of this result, functional class, speed limit, access control, and ADT will have to be considered simultaneously. Therefore, more detailed information, such as that mentioned above, will be required.

AUTHORS' CLOSURE

We very much appreciate the interest of the discussant and quite a number of other reviewers of our paper examining the relationship of median width and highway accident rates. Obviously this is a subject area of considerable interest.

The first issue raised by the discussant was that "it is entirely possible that the paved inside shoulder width was associated with the accident rate, not the rest of median width." As pointed out, we did not separate the inside shoulder width from the remainder of the median width in the analyses, primarily because of the difficulty of determining where the median/shoulder "begins" for unpaved shoulders (approximately 43 percent of the data). Although it is an interesting hypothesis, we continue to believe that the effect seen is from the total median width rather than just the paved shoulders. Unfortunately, we are not able to reanalyze the data at this time.

After the question was raised, we reexamined the available Illinois roadlog file. (Utah data were unavailable at this time.) In the first place, as noted above, nearly half of the sections in the study file (43 percent) were not paved (i.e., earth, sod, aggregate, surface treated, or no shoulder). Of those that were paved, virtually all were 8 ft (2.5 m) or less and most often (54 percent of the time) were found on roads with median widths of 64 ft (19.6 m) or greater. Less than 10 percent of the sections with paved inside shoulders had median widths of less than 40 ft (12.3 m), where we also saw significant effects of total width.

In short, we find it hard to imagine in this case that paved inside shoulders could account for the effects found in the analysis. However, it is an interesting hypothesis that could be explored further.

With regard to the question about right shoulder widths, only 3 of 982 Utah sections had right shoulder widths exceeding 15 ft (4.6 m) and none of 1,481 Illinois sections had right shoulder widths exceeding 13 ft (4.0 m).

With regard to the comment that "the estimated coefficients for ADT have an algebraic sign contrary to expectation," we do not see why the result that sections of freeways with higher ADTs have lower accident rates is not acceptable. Whereas lower accident frequencies would not be expected, lower accident rates may be. Is it not conceivable that sections with a higher ADT may have slower traffic speeds due to congestion, for example? It should be noted that the number of lanes is the same for all sections in this analysis, meaning higher ADT sections are more congested by definition.

The discussant notes that "the estimated regression coefficients for 'median width' have very low *t*-statistics, indicating that the effect of median width on accident rate was poorly determined from the data." The individual *t*-statistics for the median width terms are not especially relevant to whether median width has an effect. Overall, the effect of median width is significant. However, we agree that we could be more sure of the shape of the trend if the individual coefficients were significant as well. Note that median width was examined in greater detail (i.e., quadratic, cubic, and quartic functions) than other variables because it was the primary variable under investigation in this study.

Finally, the problem of collinearity is discussed in the section on statistical methods starting with the paragraph that begins with "Many of the variables included in the regression model were correlated with median width. . ." The available data do not allow clear resolution of the problem. Interactions representing the simultaneous effect of two variables at a time

were investigated, but, as stated in the paper, none were found to be significant. Simultaneously cross-classifying sections by median width, functional class, speed limit, access control and ADT is not practical because there would be too few sections in each cell of such a cross-classification. The regression approach adopted appears to be the only practical method of adjusting for other variables in this case.

Again, we appreciate these thoughtful comments and suggestions by the discussant and others and believe that the paper has generally addressed them to the extent practicable with the available data.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.

Estimation of Safety at Two-Way Stop-Controlled Intersections on Rural Highways

JAMES A. BONNESON AND PATRICK T. MCCOY

The application of the generalized linear modeling approach to the development of a model relating unsignalized intersection traffic demands to accident frequency is described. Several techniques for assessing model fit have been described and any inherent limitations noted. The model was based on the product of the intersection traffic demands raised to a power. This model form was found to explain a large portion of the variability in accidents among intersections of similar geometry and traffic control. The analysis of accident data for 125 two-way stop-controlled intersections supports the theory that the distribution of accident counts can be described by the negative binomial distribution. Also supported is the assertion that the mean accident frequency for the group of similar intersections is gamma distributed. Knowledge of these distributions and their parametric values can be used to identify hazardous locations and the true effect of safety treatments on accident frequency.

The intersection of a minor roadway with a rural highway presents several safety problems. Accident rates may be higher at intersections on rural highways than at intersections on urban highways. This trend may be attributed to the higher speeds on the rural highway and a reduced driver expectancy for intersecting traffic movements on rural highways. The combination of high-speed rural highway traffic with low-speed intersecting traffic (i.e., crossing or turning vehicles) can lead to significant speed differentials and an increased potential for accidents. If the highway has a multilane (or expressway) cross section with wide median, the intersection will have a large conflict area. This characteristic, combined with high expressway speeds, can further degrade the safety of the crossing or turning maneuver.

The frequency of access points, the amount of roadway lighting, and the magnitude of traffic demands are typically lower in rural than in urban areas. These characteristics tend to make drivers on rural highways more relaxed and potentially less attentive. As a result, the highway driver's expectancy for turning or crossing vehicles may be relatively low, which can also increase accident potential at intersections on rural highways.

Unfortunately, little is known about the safety of unsignalized intersections on rural highways. As a result, it is difficult to accurately identify intersections that are truly hazardous. Moreover, it is difficult to fully assess the safety benefits of any corrective measures (e.g., advance signing, signalization) applied to hazardous intersections.

Civil Engineering Department, University of Nebraska-Lincoln, Lincoln, Nebr. 68588-0531.

The objective of this research was to develop a methodology for assessing the safety and efficiency of both intersections and interchanges on rural highways. This paper describes the development of a safety-prediction model for rural intersections. The safety of an intersection is defined as the expected number of accidents per year. The research described is part of a more comprehensive analysis of the economic benefits and costs of interchanges, relative to intersections, on rural expressways conducted by the authors for the Nebraska Department of Roads (*1*).

MODELING CONSIDERATIONS

Many factors affect the number of accidents that occur at an intersection. The factors can be categorized into those representing exposure to potential accident events and those affecting the probability that a given potential accident event will result in an accident. Factors that represent exposure include time period and traffic demand. Factors that affect accident probability include urban/rural environment, traffic control, frequency of access points, speed limit, shoulder width, median type, median width, lighting level, availability of left-turn bays, number of legs, and number of traffic lanes.

Several modeling issues were considered in determining the form of the safety-prediction model and the types of data needed to calibrate it. One of the more critical issues is sample size. The high variability of accident data tends to translate into the need for large sample sizes to establish sound statistical relationships. One way to overcome the uncertainties of high variability is to increase the size of the data base. An increase could be accomplished by including more intersections in the data base (if available) or by increasing the duration of the time period that brackets the data (provided that environmental or driver behavior changes do not become significant).

Similarity among intersections in the data base is another issue. The calibrated model must be able to predict accident frequency for the "typical" type of intersection being considered. Thus, a large data base containing many different types of intersections should be subset to yield a smaller data base having intersections more consistent with the type of intersection being considered. For example, in this research, we initially subset the statewide data base to remove all urban intersections.

Although subsetting has the beneficial effect of increasing similarity in the resultant data base, it is achieved by elimi-

nating a portion of the data base. Thus, it has the disadvantage of reducing the available sample size. The trade-offs between sample size and sample similarity must be carefully considered and a proper balance achieved. In general, the amount of subsetting is generally limited to that which will allow the questions of the research to be answered without compromising the analyst's predetermined sample size requirement.

LITERATURE REVIEW

A comprehensive review of the literature regarding accident frequency prediction models before 1981 has been provided by Satterthwaite (*2*). A more recent review has been provided by Hauer et al. (*3*). In the latter work, the authors cite considerable evidence supporting the "product-of-flows-to-power" model, wherein the expected accident frequency is a function of the product of traffic demands entering the junction. In most instances, the traffic demands are raised to a power less than unity, indicating a nonlinear relationship between demand and accident frequency. Some researchers [e.g., Hauer et al. (*3*)] have considered only the flows for the conflicting traffic patterns. Others [e.g., Van Every (*4*)] relate expected accident frequency to the product of the average daily traffic demands on the major and minor roadways at the junction. The latter approach was particularly appealing because it does not require detailed information about travel patterns through the intersection. The model used by Van Every (*4*) is shown in Equation 1.

$$A = 0.00042 T_m^{0.5} T_c^{0.5} \quad (1)$$

where

A = expected annual accident frequency,
T_m = major road traffic demand (veh/day), and
T_c = minor (cross) road traffic demand (veh/day).

On the basis of this review of the literature, it was determined that both exposure measures (i.e., time period and traffic demand) were essential components of the data base. It was also determined that environment (urban versus rural), traffic control (signal, sign), and geometry (number of legs) were among the most important factors to be considered when subsetting the accident data base.

MODEL DEVELOPMENT

Data Base

The data base used for this study was obtained from FHWA via the Highway Safety Research Center at the University of North Carolina. This data base was a subset of FHWA's Highway Safety Information System (HSIS) (*5*). HSIS integrates the accident, roadway design, and traffic volume data from the highway departments of Utah, Minnesota, Illinois, Maine, and Michigan. The inclusion of roadway design and average daily traffic volume (ADT) data in the accident data base is one of HSIS's key features; it is a feature not found in most state accident record systems. This type of comprehensive data base was essential to the calibration of a safety-prediction model because it allowed the relationship between geometry, traffic demand, and accidents to be fully explored.

The data base includes all accidents during 1985, 1986, and 1987 for Minnesota. The other states in HSIS were not included because they did not recognize junctions as entities or explicitly differentiate between intersections and interchanges.

The data base can be described as a junction-based file because all data subsetting is related to the type of junction. The subset factors were selected such that the data base included intersections that are similar with regard to nonurban environment, four legs only, two-way stop-control, and intersection geometry (i.e., not an interchange ramp-junction). Once the appropriate intersections were identified, traffic demand and accident data files were scanned to find the corresponding demands and accidents (if any) for these locations. Accidents were assumed to be intersection related if they occurred within 153 m (500 ft) of the junction.

The resulting data base contained major and minor road ADTs for 125 intersections, which experienced 250 accidents in the 3-year period. Further subsetting based on accident pattern, number of lanes, median width, and so forth was not considered because the sample size was determined to be at its smallest acceptable value. A summary of the data base used in this study is given in Table 1.

The data base included two general types of major road geometry. Included were 108 intersections with a two-lane major road and no median and 17 intersections with a four-lane major road and a median of 10.4 m (34 ft) or more. The two-lane roads generally had major road ADTs under 4,000,

TABLE 1 Accident Data Base Statistics[1]

Variable	Statistic			
	Minimum	Maximum	Mean	Std. Dev.
Major Road ADT	430	37,900	4,030	6,140
Minor Road ADT	45	8,850	680	1,060
Accidents/year[2]	0	7	0.67	1.20

[1]Statistics based on 125 rural highway intersections with two-way stop-control.

[2]Average annual accident frequency (based on data for three years).

whereas the four-lane roads had ADTs over 6,000; there was some overlap in the ADT range of 4,000 to 6,000.

This dichotomy in the data base was considered desirable because it related to the objectives of the research. In this regard, the question posed by the research was, What are the expected number of accidents at a given highway junction if an unsignalized intersection or an interchange were built there? The answer would be used to make "planning-level" decisions regarding the most appropriate junction type. To answer this question, safety models were developed for each junction type (only the intersection model is described herein) for a common range of traffic demands. With the dichotomized data base, we were able to analyze intersections with ADTs well into the range found at interchanges. In application, it is understood that the model's predictions would reflect the likelihood that a two-lane major road exists for lower-demand situations and a four-lane road exists for higher-demand situations.

The relationship between average annual accident frequency and traffic demand is given in Table 2 for the two-way stop-controlled intersections included in this study. Examination of the row and column summaries indicates a positive correlation between traffic demand and average annual accident frequency. The traffic demand ranges in Table 2 were selected such that approximately one-fifth of the total number of accidents are located in each row and column. The intent of distributing the accidents in this manner is to obtain an equal weight, in terms of observations, underlying the average annual accident frequency provided in the row and column summaries.

Modeling Techniques

Model Structure

The nonlinear relationship between accident frequency and traffic demand evidenced in Table 2 is consistent with the nonlinear product-of-flows-to-power formulation, as advocated by others (*2–4*). As a result, the following model form was considered:

$$E(m) = b_0 T_m^{b_1} T_c^{b_2} \tag{2}$$

where

$E(m)$ = expected accident frequency,
b_i = regression constants,
T_m = major road traffic demand (veh/day), and
T_c = minor (cross) road traffic demand (veh/day).

The approach taken in calibrating the accident prediction model was based on procedures described by Hauer et al. (*3*), who have argued against using traditional least-squares regression of accident data because of violations in two assumptions (i.e., normally distributed error structure and constant variance) on which this type of analysis is based. Instead, Hauer et al. (*3*) advocate the use of a generalized linear model [e.g., GLIM (*6*)] wherein these assumptions are avoided, thereby yielding a better predictor of accident frequency as influenced by other factors.

Before proceeding it is important to define the safety m of an intersection as its mean accident frequency. This quantity can be estimated by taking the average of the m's [$E(m)$] for a large number of similar intersections, each having identical traffic demands. In this context, similar intersections have one or more geometric or traffic control characteristics in common. The estimate becomes more stable as the intersections become more similar (i.e., as the number of characteristics that they have in common increases).

In the last few years, Hauer et al. (*3*) and others have convincingly argued that the distribution of accident counts for a group of similar sites (e.g., intersections, road sections) can be described by the family of compound Poisson distributions. In this context, there are two sources of variability underlying the count distribution. One source stems from the

TABLE 2 Cross-Tabulation of Accidents by Daily Traffic Demand for Two-Way Stop-Controlled Intersections

Major Road (veh/day)	Minor Road (veh/day)					Row Summary
	45-517	518-1070	1071-1682	1683-3137	3138-8850	
430-2037	0.17[a] 9.67/57[b]	0.81 5.67/7	1.00 1/1	0 0/0	0 0/0	0.25 16.33/65
2038-3887	0.17 2.66/16	0.61 3.67/6	1.13 5.67/5	1.67 1.67/1	2.67 2.67/1	0.56 16.33/29
3888-7675	0.46 3.67/8	1.50 3/2	1.84 3.67/2	1.17 2.33/2	5.00 5/1	1.18 17.67/15
7676-17150	0 0/0	1.00 5/5	3.34 6.67/2	0 0/0	4.00 8/2	2.19 19.67/9
17151-37900	0 0/0	1.33 1.33/1	0.67 1.33/2	2.67 10.67/4	0 0/0	1.90 13.33/7
Column Summary	0.20 16.0/81	0.89 18.67/21	1.53 18.33/12	2.10 14.67/7	3.92 15.67/4	0.67 83.33/125

[a]The top number in each cell represents the average annual accident frequency for the corresponding range of major and minor daily traffic demands (based on data for three years).

[b]The bottom numbers in each cell represent the average annual number of accidents and number of junctions (annual accidents / total junctions) for the range of daily traffic demands.

differences in the m's among the similar sites. The other stems from the randomness in accident frequency at any given site, which is traditionally described as Poisson.

In spite of being similar, each site has its own regional character and driver population, giving it a unique mean accident frequency, m. Thus, the distribution of m's within the group of similar sites can be described by a probability density function with mean $E(m)$ and variance $V(m)$. Hauer et al. (*3*) have shown this distribution to be adequately described by the gamma density function.

Abbess et al. (*7*) have shown that if accident occurrence at a particular site is Poisson distributed, the distribution of accidents around the $E(m)$ of a group of sites can be described by the negative binomial distribution. The variance of this distribution is

$$V(x) = E(m) + \frac{E(m)^2}{k} \tag{3}$$

where x is the observed accident count. Since the variance of the Poisson distribution is $E(m)$, it is apparent that the variance of the negative binomial distribution exceeds that of the Poisson by $[E(m)^2]/k$. Hauer et al. (*3*) have shown that this latter quantity is equivalent to the variance of the mean accident frequency for the group of similar sites, $V(m)$. Hauer et al. (*3*) have also shown that the parameter k can be estimated by fitting Equation 3 to $V(x)$ and $E(m)$ estimates for the group of similar sites. The $V(x)$ is estimated as the squared difference between the accident count and the corresponding $E(m)$ for each site in the group.

Generalized Linear Model

The analysis tool used to estimate the model coefficients was the nonlinear regression procedure (NLIN) in the SAS statistical software package (*8*). This procedure is sufficiently general that it can be modified to accommodate error structures that are not normally distributed. It can also be easily modified to yield maximum-likelihood model coefficients. With these modifications, the NLIN procedure can be used as a generalized linear model similar to the GLIM package (*6*). An example application of NLIN to generalized linear modeling is described in the SAS documentation (*8*, Example 6). The SAS code described in this documentation was modified (due to some errors in printing) and enhanced to include the negative binomial and gamma distributions.

The generalized linear modeling approach relates a linear predictive equation to the expected value of an observation via a "link function." The link function equates this linear predictive relationship to a nonlinear, and perhaps bounded, dependent variable. One link function is theoretically related to the error structure of the data. This link function is sometimes referred to as the "natural" (or canonical) link. The selection of the appropriate link function is often based on the distribution of the error structure; however, as noted by McCullagh and Nelder (*9*), this is not a requirement. The natural link functions for the Poisson distribution and negative binomial distributions are given by Equations 4 and 5, respectively:

$$\eta = \ln[E(m)] \tag{4}$$

$$\eta = \ln\left[\frac{E(m)}{k + E(m)}\right] \tag{5}$$

where the linear predictive equation is

$$\eta = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n \tag{6}$$

To implement proposed model form, it was necessary to take the inverse of the link function [i.e., $E(m) = f^{-1}(\eta)$], equate it to the right-hand side of Equation 2, and solve for η. For the Poisson link function, the resulting linear predictive model takes the following form:

$$\eta = \ln(n) + \ln(b_0) + b_1\ln(T_m) + b_2\ln(T_c) \tag{7}$$

where $\ln(n)$ is termed the offset variable with an implied coefficient of 1.0. For this type of analysis, the offset variable is equivalent to the number of years underlying the observed count (in this study, $n = 3$ years for all observations). This linear predictive model form lends itself to further expansion if additional regression parameters are desired in the model. For example, a study by Pickering et al. (*10*), where additional parameters were included to examine the effects of various geometric elements on accident frequency at unsignalized T-intersections, illustrates the use of this model.

A similar calculation of the linear predictive model form was not as successful for the negative binomial link function. In fact, it was not possible to implement the proposed model (i.e., Equation 2) in its intended form using the natural link for the negative binomial structure. Because of this loss of generality, and in recognition that it is not a requirement to use the natural link, the Poisson link was used for all analyses in this study.

Quality of Fit

Several statistics are available from the NLIN procedure for assessing the model fit and the significance of model coefficients. One measure of model fit provided by NLIN is the generalized Pearson X^2 statistic. This statistic is calculated as

$$X^2 = \sum \frac{[x - \hat{E}(m)]^2}{\hat{V}(x)} \tag{8}$$

where $\hat{V}(x)$ is estimated from Equation 3 by substituting $\hat{E}(m)$ for $E(m)$. This statistic is available from NLIN as the "Weighted Sum of Squares" for the Residual. McCullagh and Nelder (*9*) indicate that this statistic follows the χ^2 distribution with $n - p - 1$ degrees of freedom, where n is the number of observations and p is the number of model parameters. This statistic is asymptotic to the χ^2 distribution for larger sample sizes and exact for normally distributed error structures. As noted by McCullagh and Nelder (*9*), this statistic is not well defined in terms of minimum sample size when applied to nonnormal distributions; therefore, it probably should not be used as an absolute measure of model significance.

Another, more subjective, measure of model fit can be obtained from a plot of the prediction ratio versus the estimate

of the expected accident frequency [i.e., $\hat{E}(m)$]. In this context, the prediction ratio is the ratio of the observed accident frequency to the expected accident frequency. This type of plot yields a visual assessment of the predictive capability of the model over the full range of $\hat{E}(m)$. A well-fitted model would have the prediction ratios symmetric about 1.0 over the range of $\hat{E}(m)$. This technique was applied by Hauer et al. (*3*) in a recent study of safety at signalized intersections.

The significance of the parameter coefficients (with respect to the hypothesis that they equal zero) is also helpful in assessing the relevance of model factors. In this regard, NLIN provides the standard error and 95 percent confidence interval for each coefficient. Because the Pearson X^2 statistic (i.e., Equation 8) has some limitations, the significance of the individual parameter coefficients may represent a more realistic measure of model fit.

Finally, the dispersion parameter σ_d is noted by McCullagh and Nelder (*9*) to be a useful statistic for assessing the amount of variation in the observed data. This statistic can be calculated by dividing Equation 8 by the quantity $n - p$. It is also available from NLIN as the "Weighted Mean Square" for the Residual. A dispersion parameter near 1.0 indicates that the assumed error structure is approximately equivalent to that found in the data. For example, if a Poisson error structure is assumed [i.e., $V(x) = E(m)$] and the dispersion parameter is 1.68, the data have greater dispersion than is explained by the Poisson distribution. In this situation, the negative binomial distribution might be considered, since it has a larger variance than does the Poisson (see Equation 3).

Analysis Procedure

Coefficient estimation for the proposed model was a multistep process. First, the data were analyzed using a Poisson error structure. Then NLIN was used to fit Equation 3 to the squared residuals from the first analysis [Hauer et al. (*3*) indicate that the squared residuals can be used as an estimate of $V(x)$]. This second step yielded an estimate of the k parameter and a measure of its statistical significance. The need for a third analysis step was based on an assessment of the dispersion parameter and the k parameter significance. If the dispersion parameter was more than 1.0 and the k parameter was statistically significant, a third analysis step was conducted using the negative binomial error structure with k from Step 2 as an estimate of the shape parameter. The residuals from this analysis were analyzed in a fourth step to determine a new k parameter. The third and fourth steps were repeated until convergence on a value of k. This procedure is consistent with that described by Hauer et al. (*3*).

Calibrated Models

The results of the first analysis step (Poisson error) are given in Table 3. As indicated in the table, there was sufficient dispersion (i.e., $\sigma_d > 1.0$) to justify further analysis in terms of the third and fourth steps (negative binomial error). The second step indicated that a k parameter of 3.9 was significant. The Pearson X^2 statistic indicated that the assumed Poisson error structure did not account for a significant portion of the dispersion in the observations. Thus, further analysis via Steps 3 and 4 appeared warranted.

After two iterations of Steps 3 and 4 using a negative binomial error structure, the k parameter converged to 4.0. The fit of Equation 3 with the squared residuals is shown in Figure 1. Each circle represents the expected accident frequency averaged for several intersections. These intersections were ranked in ascending order of accident frequency before averaging. Average values eliminate the visual "noise" associated with a plot of 125 individual data points and thereby better illustrate the underlying curvilinear trend. The number of intersections included in each average was based on the desire to plot 9 or 10 points that have weight as nearly equal as possible (8 points at a weight of 14 intersections and 1 at a weight of 13 worked best). Although most of the observations are clustered near the origin, there appears to be a definite correlation with the variance function.

On the basis of this analysis, the following model was developed for predicting the annual expected accident frequency for two-way stop-controlled intersections on rural highways:

$$E(m) = 0.692\left(\frac{T_m}{1{,}000}\right)^{0.256}\left(\frac{T_c}{1{,}000}\right)^{0.831} \tag{9}$$

The coefficients in this model are significant at a 95 percent level of confidence (i.e., 5 percent chance of false rejection). The Pearson x^2 statistic was significant, indicating that the

TABLE 3 Model Statistics

Statistic		Analysis Steps 1-2	Analysis Steps 3-4
Model	Error Distribution	Poisson	Neg. Binomial (k=4.0)
	Pearson X^2	215	136*
	$\chi^2_{0.05,n-p-1}$ (n = 125, p = 3)	147	147
	Dispersion Parameter, σ_d	1.76	1.11
Parameter[1]	Intercept, b_0	0.650 (0.080)*	0.692 (0.121)*
	Major Road, b_1	0.292 (0.065)*	0.256 (0.099)*
	Minor Road, b_2	0.791 (0.066)*	0.831 (0.102)*

* Denotes significance at a 95-percent confidence level (5 percent chance of false rejection).

[1] Parameter values include coefficient estimates and standard error. Standard error is in parentheses.

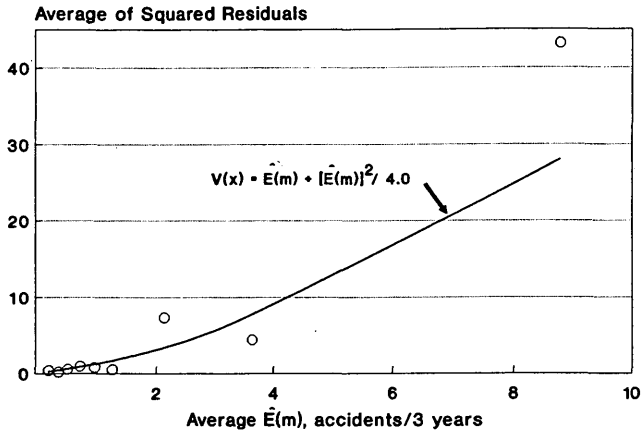


FIGURE 1 Relationship between expected accident frequency and estimated accident variance.

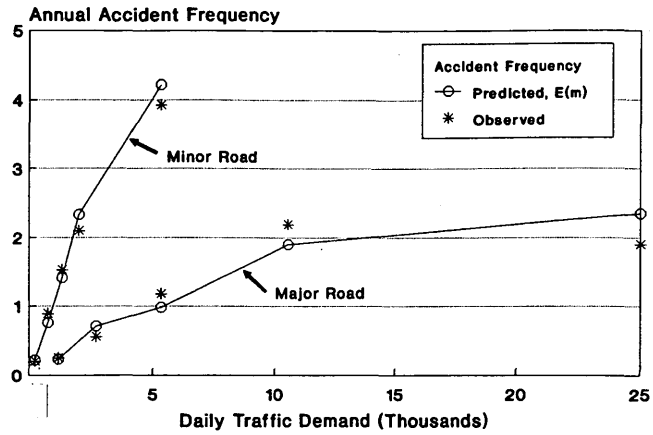


FIGURE 3 Accident frequency as a function of daily traffic demand.

negative binomial error structure is able to explain a significant portion of the deviation from the predicted values. The dispersion parameter was much nearer to 1.0 than for the first analysis, which further supports the selection of the negative binomial structure.

The fit of the model to the data can be assessed using the open circles in Figure 2. The ratio of the observed accident count to the expected accident frequency for a 3-year period is shown. The desired symmetry of observations around 1.0 is not as apparent as desired; however, the large black dots (averages of 14 ratios previously ranked in ascending order) indicate that a symmetry around 1.0 exists.

The curvilinear trend suggested by some combinations of the open circles stems from the reciprocal nature of the plotted quantities [i.e., $E(m)$ versus $1/E(m)$]. For example, the most distinct curvilinear combination in Figure 2 represents those intersections having an observed accident count of one accident per 3 years.

The predictive capability of the model is shown in Figure 3. The starred data points represent the average annual accident frequencies from the row and column summaries of Table 2. Each open circle represents the average of the predicted annual accident frequency for the intersections that are included in the starred data points. The number of intersections that underlies each pair of data points is also provided in the row and column summaries of Table 2.

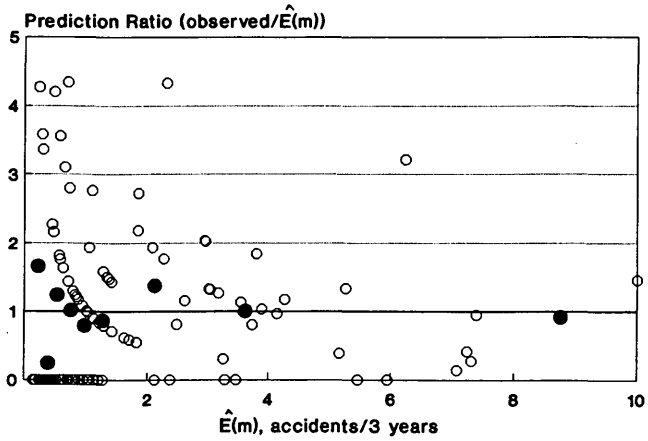


FIGURE 2 Evaluation of model fit using prediction ratios.

Figure 3 provides further support for the exponential relationship between accident frequency and traffic demand. The close agreement between the observed and predicted frequencies suggests that Equation 9 is a good predictor of the expected number of accidents at a typical rural, two-way stop-controlled intersection. A comparison of Equation 9 with the model proposed by Van Every (*4*) (i.e., Equation 1) indicated good agreement between the models.

CONCLUSIONS

This paper describes the application of the generalized linear modeling approach to the development of a model relating intersection traffic demands to accident frequency. The general linear model was implemented using the nonlinear regression procedure (NLIN) of the SAS program (*8*) with appropriate modification.

Several techniques for assessing model fit have been described and any inherent limitations noted. Of the techniques, the authors place the most confidence in the plot of prediction ratio versus expected number of accidents. This plot indicates the amount of dispersion in the predicted values as well as the existence of any model bias over the range of accident frequencies considered.

The product-of-flows-to-power model formulation appears to explain a large portion of the variability in accidents among intersections of similar geometry and traffic control. The strength of this model format has been shown by others (*2–4,10*) and in this paper as applied to two-way stop-controlled intersections on rural highways. The form of this model suggests that mean accident frequency increases in a nonlinear fashion with increasing major or minor road demand.

The analysis of accident data for 125 two-way stop-controlled intersections supports the theory that the distribution of accident counts can be described by the negative binomial distribution. Also supported is the assertion that the mean accident frequency for the group of similar intersections is gamma distributed. Knowledge of these distributions and their parametric values can be used to identify hazardous locations

and the true effect of safety treatments on accident frequency. A considerable amount of research in this area has been performed by Hauer and Persaud (*11*).

ACKNOWLEDGMENTS

The authors wish to express their appreciation to the individuals and agencies who assisted in making this project successful. These individuals include Duane Eitel of the Nebraska Department of Roads, Jeffrey F. Paniati of FHWA, and Forrest M. Council of the University of North Carolina. A special thank-you is extended to the latter two gentlemen for making the HSIS data base available. The authors would also like to acknowledge the technical assistance of Brian Moen and Jim Kollbaum, students at the University of Nebraska-Lincoln.

REFERENCES

1. J. A. Bonneson and P. T. McCoy. *Interchange vs. At-Grade Intersection on Rural Expressways.* Research Report TRP-02-25-91. Nebraska Department of Roads, Lincoln, May 1992.
2. S. P. Satterthwaite. *A Survey of Research into Relationships Between Traffic Accidents and Traffic Volumes.* TRRL SP 692. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1981.
3. E. Hauer, J. C. N. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 48–61.
4. B. E. Van Every. A Guide to the Economic Justification of Rural Grade Separations. *Australian Road Research*, Vol. 12, No. 3, pp. 147–154.
5. F. M. Council and J. F. Paniati. The Highway Safety Information System. *Public Roads, A Journal of Highway Research and Development*, Vol. 54, No. 3, Dec. 1990, pp. 234–239.
6. R. J. Baker and J. A. Nelder. *The GLIM System: Release 3, Manual.* Rothhamsted Experimental Station, Harpenden, England, 1978.
7. C. Abbess, D. Jarrett, and C. C. Wright. Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the "Regression-to-Mean" Effect. *Traffic Engineering and Control*, Vol. 22, No. 10, Oct. 1981, pp. 535–542.
8. *SAS/STAT User's Guide, Version 6* (fourth edition, Vol. 2, Chapter 29). SAS Institute, Inc., Cary, N.C., 1990.
9. P. McCullagh and J. A. Nelder. *Generalized Linear Models.* Chapman and Hall, New York, 1983.
10. D. Pickering, R. D. Hall, and M. Grimmer. *Accidents at Rural T-Junctions.* TRRL Research Report 65. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1986.
11. E. Hauer and B. N. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1987, pp. 131–140.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Investigation of the Effectiveness of the Virginia Habitual Offender Act

CHERYL W. LYNN, JACK D. JERNIGAN, ANTHONY NORRIS, AND PATRICIA FRONING

In 1968, Virginia passed the Habitual Offender Act (the Act), one of the first laws in the United States directed at motorists who repeatedly violate traffic laws. Persons convicted as habitual offenders are subject to long-term license revocation, and those who violate this revocation may be incarcerated. A study was conducted to determine the effectiveness of the Act in enhancing traffic safety in Virginia. A sizable number of individuals whose driving records include a sufficient number of convictions to allow the Virginia Department of Motor Vehicles to certify them as habitual offenders are never brought before the courts on the charge. The existence of such a group of certified yet nonadjudicated habitual offenders is an indication that the procedures implementing the Act need to be changed to promote a more global implementation. However, the existence of this group allowed the researchers to compare a sample of certified habitual offenders with a group that had been adjudicated. In general, the adjudicated group had more prior convictions for driving under the influence of drugs or alcohol, and the certified group had more convictions for operating under a suspended operator's license and more convictions that were defined as minor offenses under the Act. However, the adjudicated group had fewer subsequent traffic convictions and crashes and were conviction-free and crash-free for a longer period of time. These data indicate that adjudication under the Act may enhance traffic safety.

A number of states have enacted statutes directed at motorists who repeatedly violate traffic laws. These "habitual offender" statutes seek to provide maximum safety for all drivers by denying the privilege of driving to persons convicted of a specified number and type of traffic offenses. Persons convicted as habitual offenders are subject to long-term license revocation, and those who violate this revocation may be incarcerated.

In 1968, Virginia enacted one of the first habitual offender laws in the United States, preceded only by Colorado in 1953 and Delaware in 1958. The Virginia Traffic Safety Study Commission recommended the legislation in a report to the governor and the general assembly in 1967, stating its belief that there were many serious offenses that warranted permanent revocation of driving privileges. The commission endorsed and recommended the passage of a modified version of a habitual offender bill that had been drafted by the Virginia Association of Insurance Agents, Inc.

Since the enactment of the Habitual Offender Act (the Act) more than 20 years ago, there have been no published studies on its effectiveness in promoting traffic safety. Likewise, there has been no analysis of whether the sanctions imposed by the Act have accomplished the objective of reducing the number of crashes and convictions of persons adjudicated as habitual offenders. Because of the lack of information concerning the Act's effectiveness, some members of the Advisory Committee to the Commission on VASAP urged the committee's Subcommittee on Habitual Offenders to conduct a review of the Act. The subcommittee requested that the Virginia Transportation Research Council (VTRC) study how the Act has affected traffic safety.

The Act defines a habitual offender as any resident or nonresident whose driving record, as maintained by the Virginia Department of Motor Vehicles (DMV), shows an accumulation of 3 major offenses, 12 minor offenses, or a total of 12 major and minor offenses, all within a 10-year period. Major offenses include voluntary or involuntary manslaughter resulting from the operation of a motor vehicle; driving while under the influence of drugs or alcohol (DUI); driving on a suspended or revoked license (SOL); perjury as to matters pertaining to the motor vehicle laws; any felony involving the motor vehicle laws or the use of a motor vehicle; and hit and run involving injury, death, or property damage in excess of $500. Minor offenses under the Act are those that require the DMV or authorize a court to suspend or revoke a driver's license for a period of 30 days or more. The court does not have to actually suspend or revoke a license for an offense to be counted toward habitual offender certification.

Out-of-state convictions and convictions under local Virginia ordinances that substantially conform to the offenses listed in the Act are included for habitual offender status. Multiple offenses committed in a 6-hr period are counted as one offense provided a driver has no prior chargeable violations. Once a driver has one or more chargeable convictions, all future convictions are counted separately regardless of the time period in which they occur.

Once a driver's record has been identified by DMV as qualifying under the Act, DMV must certify three abstracts of the convictions that counted toward the habitual offender certification to the commonwealth's attorney of the political subdivision in which the person resides. In the case of a nonresident, the commonwealth's attorney of Richmond is sent the three abstracts. The abstract is prima facie evidence that the person was duly convicted. A person who denies any of the convictions on the abstract has the burden of proving that the questioned information is not correct.

The commonwealth's attorney then has the discretion to pursue one of the following courses of action:

1. To file an information against the certified driver in a court of jurisdiction (an information is an official criminal

Virginia Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22903-0817.

charge presented by the commonwealth's attorney without the interposition of a grand jury),

2. Not to file an information, or
3. To ignore the certification.

If the court finds that the accused is not the person named in the abstract or that the individual is not a habitual offender under the terms of the Act, the proceedings are dismissed and DMV is notified of these results. If the person is found to be a habitual offender, the court directs the person to surrender to the court his or her license to drive a motor vehicle. The court further orders the offender not to drive on the commonwealth's highways.

In any case where the accused is charged with SOL, the Act directs the court to determine whether the person is currently under a habitual offender revocation. If the court finds that the accused has been held to be a habitual offender, it certifies the case to a court of record for trial. Any person who is under a habitual offender revocation and is subsequently convicted of violating that order is confined to a state correctional facility for not less than 1 and not more than 5 years or confined in jail for 12 months. No part of the sentence may be suspended except any portion in excess of 1 year or where the accused drove in the case of an extreme emergency to save life or limb.

Although the court order is for a permanent revocation, the offender may petition the court for reinstatement after a 10-year period. The burden of persuasion is on the petitioner to show good cause why the revocation should be lifted. The court, at its discretion, may restore the person's driving privileges under whatever conditions it prescribes.

There are three exceptions under which the 10-year period may be shortened. The first is for individuals who were adjudged to be habitual offenders in part on the basis of findings of "not innocent" as juveniles. The offender may petition the court for a return of driving privileges after turning 18 years old.

The second exception is for individuals adjudged habitual offenders in part on the basis of convictions for DUI. Upon petition, a court may issue a restricted license after 3 years and a full license after 5 years provided that (*a*) the petitioner was psychologically dependent on or addicted to alcohol or drugs at the time of the previous conviction, (*b*) the petitioner is not addicted to or psychologically dependent on alcohol or drugs at the time of the hearing, and (*c*) the petitioner is no longer perceived as a threat to himself or herself or the public while operating a motor vehicle.

The final exception is for individuals who were found to be habitual offenders in part on the basis of convictions for SOL where the suspension or revocation was due to either a failure to pay fines; furnish proof of financial responsibility; or satisfy a judgment, provided the judgment has been paid before the petition is filed. These individuals may petition after a 5-year period and may have their driving privileges reinstated provided the court determines that the petitioner is no threat to himself or herself or others while operating a motor vehicle.

PURPOSE AND SCOPE

The purpose of this study was to gain insight into how habitual offender acts work and how Virginia's act has affected traffic safety. The study further sought to determine what types of offenses typically result in an individual being certified a habitual offender and whether certain types of offenses are more likely to result in an individual being later adjudged as a habitual offender.

METHOD

How Habitual Offender Acts Work

Studies investigating the effectiveness of habitual offender programs in other states were critically reviewed and summarized. Next, a survey of other state statutes was conducted, and statutes dealing with the most serious repeat offenders were described.

Impact of Virginia's Act on Traffic Safety

One way to measure the effectiveness of the Act on traffic safety would be to measure how many of those drivers who have their privilege to drive revoked no longer operate a motor vehicle. However, determining whether habitual offenders still drive is both methodologically and practically impossible. Instead, the performance measure chosen for this study was the effectiveness of the Act in reducing future traffic crashes and convictions—its ultimate intent. Thus, even if persons adjudicated under the Act continued to drive, albeit illegally, the Act would be considered effective if those adjudicated became less of a traffic safety risk by driving less, driving more safely, or both.

To carry out this objective, the driving records of habitual offenders certified under the Act but not adjudicated were compared with those both certified and adjudicated. Because many habitual offenders are not brought to trial, an adequate sample of such drivers can be compared with drivers who are adjudicated by the court system and have their license revoked.

Driver history data for the samples were obtained from the internal DMV "transcript of record" printouts, which are the most comprehensive DMV driver history records. The researchers initially hoped to draw a 25 percent sample of individuals who were certified as habitual offenders by DMV in 1986. Going back to 1986 would allow individuals to be tracked for up to 5 years, during which some of the cases would eventually be adjudicated and others would not have been processed through the courts. However, this proved to be impossible because until 1992 the certification date was purged from the driver history file once a certified habitual offender was adjudicated. Hence, DMV has records only of habitual offenders certified in 1986 who have not yet been adjudicated. To rectify this problem, a 25 percent sample of those certified in 1986 who had not yet been adjudicated and a 25 percent sample of those who were adjudicated in 1986 were selected.

Once records of certified and adjudicated habitual offenders were examined, however, it became clear that it was a common practice for DMV to recertify drivers as new convictions were recorded. In addition, in some cases, these recertified drivers were then adjudicated, thus placing them in

the adjudicated group. These discrepancies in group membership could have been alleviated by restricting both the adjudicated and certified groups to drivers who had been certified only once. However, since the certification date and the record of certification are purged from the DMV record once a certified habitual offender is adjudicated, it is impossible to identify recertified drivers in the adjudicated group. Since it was impossible to remove equivalent recertified drivers from both the adjudicated and certified groups, it was decided that recertified drivers would be included in the analysis. In addition, a few drivers had been readjudicated, but because of their small number, they were left in the adjudicated group and not subjected to separate analysis. Thus, two groups of drivers were compared in this study: (*a*) adjudicated (drivers who had been certified and then adjudicated at least once) and (*b*) certified (drivers who had been certified at least once but were never adjudicated).

An obvious limitation of this sampling strategy is that the certified and adjudicated samples may not be comparable. That is, there are likely reasons why an individual might fall into one group or another. It might be the case that commonwealth's attorneys pursue adjudication against only the worst offenders or that the certified group is composed of a transient population more difficult to contact. Assignment to the groups, therefore, may not be random. Hence, it was necessary for the researchers to determine whether the two samples were comparable in terms of their previous driving record before analyzing their subsequent driving behavior. Using a t test at the 0.05 significance level, the groups were compared on the basis of five variables: age, sex, number of prior convictions for DUI, number of prior convictions for SOL, and prior convictions for minor violations under the Act. As Table 1 indicates, there were significant differences between the prior driving records of the certified and adjudicated groups. The adjudicated group had more previous DUI convictions than the certified group, but the certified group had more prior SOL and minor violation convictions. There were no significant differences between the two groups in terms of age and sex, with the average age being approximately 31 years and approximately 95 percent of each sample being male.

RESULTS

Literature Review

Very little research has been conducted with regard to the operational aspects of effectiveness of habitual offender legislation. North Carolina evaluated its 1969 Habitual Offender Law in 1975 by comparing the 2-year subsequent driving records of adjudicated habitual offenders with the subsequent records of those whose cases had been pending in the courts or had no action taken over that time period. Li and Waller (*1*) found no consistent significant differences between the two groups in terms of subsequent records. In Pennsylvania, on the other hand, Staplin (*2*) found that there was a sharp decline in traffic violations after license revocation even though 75 percent of the interviewed habitual offenders continued to drive. There are problems with both studies, however, since the North Carolina study examined only 2 years of driving behavior subsequent to revocation and the Pennsylvania study did not use a control group to document changes in subsequent driving behavior.

With regard to operational issues, a 1986 California study noted that after 2 years of implementation of a habitual offender act, with lesser penalties than Virginia's, only 4 percent of drivers eligible to be habitual offenders had been prosecuted and, of these, only 21 percent were convicted (*3*).

Survey of State Statutes

A number of states have what they call habitual offender programs, which deal with less serious offenders and fall under the purview of the driver improvement program. These programs are different from Virginia's habitual offender program and were excluded from this multistate comparison. The distinctive feature of Virginia's Act is the possibility of incarceration for a violation of the habitual offender revocation. Nineteen states other than Virginia have attempted to deal with habitual violators by including provisions for incarceration following a violation of the revocation. Table 2 gives the comparison between the Virginia statute and the habitual offender statutes of the 19 other states.

TABLE 1 Prior Convictions and Demographic Characteristics: Certified Versus Adjudicated

Variable	Mean		T	Significance
	Certified	Adjudicated		
Age	30.996 (n = 661)	31.759 (n = 611)	−1.41	N.S.
Sex	.051 (n = 662)	.044 (n = 613)	.61	N.S.
DUI	1.631 (n = 662)	1.940 (n = 613)	−4.36	$p < .01$
Minor Violations	1.100 (n = 662)	.936 (n = 613)	2.40	$p < .05$
Driving SOL	2.113 (n = 662)	1.915 (n = 613)	2.27	$p < .05$

TABLE 2 Statutory Definition of Habitual Offender Status

State	No. Major Offenses	Within No. Years	No. Minor Offenses	Within No. Years
California[a]	—	—	—	—
Colorado	3	7	10	5
Delaware	3	5	10	3
Florida	3	5	15	5
Georgia	3	5	—	—
Indiana	2 (most serious) 3 (serious)	10	10	10
Iowa	3	6	6	2
Kansas	3	5	—	—
Maine	3	5	—	—
Massachusetts	3	5	12	5
Montana[b]	*	3	—	—
New Hampshire	3	5	12	5
Oregon	3	5	20	5
Rhode Island	3	3	6	3
South Carolina	3	3	10	3
Tennessee	3	3	—	—
Vermont	8	5	—	—
Virginia	3	10	12	10
Washington	3	5	20	5
Wisconsin	4	5	12	5

[a] Major offenses for the habitual offender laws are counted only after a person has been convicted for driving on a suspended or revoked license where the revocation or suspension is based on a conviction for DUI or negligent driving. The number of qualifying offenses is counted during a 12-month period after specified offenses.

[b] Weighted offenses in point system add up to 30 points.

In defining the type of offenses used in qualifying a driver for habitual offender status, Virginia's Act differentiates between major and minor offenses. A total of 3 major offenses are needed within a 10-year period or 12 minor offenses within the same time frame (see Table 2). Twelve other states also make a distinction between major and minor offenses, and four states have no provision for minor offenses. Montana's unique system of assigning points to weight offenses includes all motor vehicle violations in one category. In California, a driver convicted of driving with a suspended or revoked license is designated a habitual traffic offender. Subsequent offenses are based on a point system with a large number of varying categories.

Like Virginia, fourteen states require three major offenses, but usually within a shorter time frame than Virginia's 10-year time period. Although Indiana's statute provides a category for three major offenses, it further designates a "most serious" category that requires only two convictions. Vermont requires eight major offenses, and Wisconsin requires four. California and Montana use a point system that varies in the number of offenses required, depending on the number of points assigned for each violation.

The number of years in which the major offenses can be accumulated varies among the states from 3 to 10. Only Virginia and Indiana extend the time period to 10 years. Colorado uses 7 years, Iowa uses 6, 11 states use 5, and the remaining 4 states use 3. Under California's system, the number of qualifying offenses are counted during the 12-month period after a specified triggering offense.

Of the 12 states that use major and minor offense categories, Virginia and 3 other states require 12 minor offenses, 2 states require 20, 1 state requires 15, 4 states require 10, and 2 states require only 6. As in the major offense category, only Virginia and Indiana extend the time period for minor offense accumulation to 10 years. Seven states use 5, three states use 3, and one state uses 2.

As can be seen in Table 3, the states are almost evenly split on the procedure followed in declaring a driver a habitual violator, with 11 states implementing an administrative process and 9 states requiring a court proceeding. The procedures used in states in which court action is taken resemble the process followed in Virginia, where the department in charge of motor vehicle records certifies a driver's record to a prosecutor, who in turn brings the action in a court proceeding. The states that require administrative action have varying procedures. Some states automatically revoke a driver's license after the threshold conviction, some provide a hearing after the revocation, and others provide an administrative hearing before the revocation.

Once a person's privilege to drive has been revoked, additional penalties are imposed if this revocation is violated. As indicated in Table 4, all states included in this analysis impose some sort of incarceration, but some states further provide for fines and an additional revocation period. Indiana

TABLE 3 Procedures for Adjudicating/Processing a Habitual Offender

State	Department Certifies/ Court Convicts	Administrative Action
California	X	—
Colorado	—	X
Delaware	X	—
Florida	—	X
Georgia	—	X
Indiana	—	X
Iowa	X	—
Kansas	X	—
Maine	—	X
Massachusetts	—	X
Montana	X	—
New Hampshire	X	—
Oregon	—	X
Rhode Island	X	—
South Carolina	—	X
Tennessee	X	—
Vermont	—	X
Virginia	X	—
Washington	—	X
Wisconsin	—	X

extends the violator's license suspension indefinitely if the initial revocation is violated. Montana extends the revocation period for an additional year. The imposition of a fine varies from $50 in Massachusetts to a possible $100,000 fine in Oregon. The differences in the length of incarceration among the states are numerous. The most lenient incarceration length, 10 days, is imposed by Massachusetts and Washington upon the first subsequent conviction. Tennessee allows for the most stringent length, a possible 6-year prison term. The typical sentence is between 1 and 5 years, as is imposed in Virginia.

Impact of Virginia's Act on Traffic Safety

The impact of the Act on traffic safety was determined by comparing the subsequent driving records of drivers who were merely certified as habitual offenders (and who may or may not be aware of this certification) with records of drivers who had been adjudicated.

Because of the statistically significant differences between the prior records of the adjudicated and certified groups, a direct comparison of subsequent driving records was not appropriate. To compare the subsequent records, an analysis of variance (controlling for significant differences in previous driving records) was used to adjust the data to make the prior driving records of the two groups equivalent. Using prior driving record as a covariate tests the independent effect of adjudication by holding factors such as number of DUIs constant for the adjudicated and certified groups.

The groups' subsequent records were compared on four variables: number of convictions for DUI, number of traffic events resulting in a conviction, number of crashes, and number of days between adjudication (or last certification date) and the date of the first traffic offense resulting in a conviction.

As can be seen in Table 5, after the prior records of the groups were statistically equated, the comparison of the subsequent driving rcords of the certified and adjudicated groups yielded the following information:

1. The group of certified drivers had more subsequent convictions for DUI and for other traffic events than did the adjudicated group.
2. The group of certified drivers had more subsequent traffic crashes than did the adjudicated group.
3. The group of certified drivers did not remain conviction-free and accident-free as long as the adjudicated group.

TABLE 4 Penalties for Driving After Being Declared a Habitual Offender

State	Revocation	Fine	Jail Term
California[a]	—	$2,000	180 days
Colorado	—	$1,000	2 years
Delaware	—	———	1–5 years
Florida	—	———	≤ 1 year
Georgia	—	$750.00	1–5 years
Indiana	Indefinite	1st offense–$10,000[b] Subsequent $10,000	1 1/2 years[b] 4 years
Iowa	—	$5,000	≤ 2 years
Kansas	—	$5,000	1–5 years
Maine	—	———	≤ 5 years
Massachusetts	—	$50–$100	≤ 10 days
Montana	1 year	$1,000	≤ 1 year
New Hampshire	—	———	1–5 years
Oregon	—	$100,000	5 years
Rhode Island	—	———	< 5 years
South Carolina	—	———	1–5 years
Tennessee	—	$1,000	1–6 years
Vermont	—	$5,000	≤ 2 years
Virginia	—	———	1–5 years
Washington	—	1st offense $500 2nd offense $500 Subsequent $500	10 days–6 months 90 days–1 year > 1 year
Wisconsin	—	$5,000	≤ 180 days

[a] Within 7 years of prior conviction.

[b] A lesser penalty of a $500 fine and a prison term of not more than 1 year can be imposed at the discretion of the court.

TABLE 5 Subsequent Offenses: Certified (N = 662) Versus Adjudicated (N = 613) Controlling for Previous Driving Under Suspension/Revocation Violations, Previous DUIs, and Previous Minor Violations

		Deviation from Grand Mean[a]		Significance Testing	
Offense	Grand Mean	Certified	Adjudicated	F[b]	Significance
DUI Convictions	.192	.04	–.05	11.56	$p < .01$
Crashes	.073	.03	–.04	18.35	$p < .01$
Traffic events	.809	.29	–.32	68.96	$p < .01$
Days to traffic event	1,116.78	–68.16	73.61	27.53	$p < .01$

[a]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[b]1 degree of freedom.

This type of post hoc analysis, although not definitive, suggests that the habitual offender statute had a positive impact on traffic safety. An alternative explanation, however, has been posed to explain these findings. As mentioned previously, persons who are adjudicated as habitual offenders and are then convicted of driving under habitual offender revocation (also referred to as the felony revoked violation) are sent to a local jail for 12 months or to a state prison for 1 to 5 years. This is not true of persons who are merely certified. It has been hypothesized that the reason adjudicated drivers were found to have fewer subsequent offenses than certified drivers is that a number of adjudicated drivers were incarcerated during the data collection period and, thus, were physically prevented from driving and incurring convictions and accidents. To test this alternative explanation, adjudicated drivers who had incurred a felony revoked conviction were removed from the analysis. As seen in Table 6, with these drivers removed, the results of this analysis still supported the finding that the Act had a positive impact on adjudicated drivers. Interestingly, when these drivers were included, the

Act seemed to have less of an impact on subsequent DUIs than when these drivers were excluded.

Impact of Virginia's Act on DUI Offenders

Adjudicated Drivers

If a habitual offender adjudication is based on at least one conviction for SOL resulting from failure to pay fines, furnish proof of financial responsibility, or satisfy a judgment, the habitual offender can petition for early restoration after 5 years. If the adjudication is based on at least one DUI, the habitual offender can petition for a restricted license after 3 years and for full restoration after 5 years. Since DMV purges information on outstanding fines and judgments 2 years after they are satisfied, it is often impossible to determine the reason a driver is convicted for SOL. Since this information is not available, very few individuals can petition for early restoration based on an SOL. Thus, the DUI exemption is responsible for most early restorations. For this reason, the impact of the Act on drivers whose previous DUI convictions contributed to their adjudication was also investigated.

Table 7 compares the number of subsequent DUIs, the number of subsequent crashes, the number of subsequent traffic events, and the time to the next traffic event for drivers whose adjudication was based on at least one DUI with drivers whose adjudication was not based on any DUIs. In addition, the records of drivers whose adjudication was based on at least three DUIs were compared with those of drivers who had no DUI. From these data, it can be seen that overall there were no significant differences in subsequent driving history between the two groups. Thus, in terms of driving behavior during the first 4 years after revocation, habitual

TABLE 6 Subsequent Offenses: Certified (N = 662) Versus Adjudicated (N = 497) Controlling for Previous Driving Under Suspension/Revocation Violations, Previous DUIs, and Previous Minor Violations (Excluding Possibly Incarcerated Offenders)

		Deviation from Grand Mean[a]		Significance Testing	
Offense	Grand Mean	Certified	Adjudicated	F[b]	Significance
DUI convictions	.169	.07	–.09	33.92	$p < .01$
Crashes	.072	.03	–.04	19.98	$p < .01$
Traffic events	.745	.35	–.47	113.84	$p < .01$
Days to traffic event	1,165.57	–116.73	155.48	104.51	$p < .01$

[a]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[b]1 degree of freedom.

TABLE 7 Subsequent Driving Records of Adjudicated Drivers With One or More DUIs (N = 503) and Those Without DUIs (N = 110)[a]

			Deviation from Grand Mean[b]		Significance Testing	
Year	Offense	Grand Mean	No DUIs	One or More DUIs	F[c]	Significance
First	DUI convictions	.038	–0.03	0.01	1.46	N.S.
	Traffic events	.135	–0.05	0.01	0.68	N.S.
Second	DUI convictions	.031	–0.03	0.01	2.33	N.S.
	Crashes	.008	–0.01	0.00	2.01	N.S.
	Traffic events	.108	–0.08	0.02	4.78	$p < .05$
Third	DUI convictions	.041	–0.02	0.00	0.97	N.S.
	Crashes	.008	–0.01	0.00	1.07	N.S.
	Traffic events	.103	–0.02	0.01	0.44	N.S.
Fourth	DUI convictions	.036	0.00	0.00	0.00	N.S.
	Crashes	.020	–0.01	0.00	0.58	N.S.
	Traffic events	.116	–0.01	0.00	0.13	N.S.
Total	DUI convictions	.145	–0.08	0.02	3.29	N.S.
	Crashes	.036	–0.03	0.01	3.06	N.S.
	Traffic events	.462	–0.17	0.04	3.39	N.S.
	Time to next event	1,201.018	90.39	–19.77	3.73	N.S.

[a]Controlling for previous driving-under-suspension/revocation violations and previous suspensions.
[b]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[c]1 degree of freedom.

offenders with one or more DUIs were similar to those with no DUI.

For the purposes of the Act, persons whose adjudication is based on one DUI combined with two other major violations are treated the same as those whose adjudication is based on at least three DUIs. No previous research or data support treating a driver with only one DUI as alcohol addicted. In addition, these data indicate that there is no rationale for treating habitual offenders with at least one previous DUI differently from those with no previous DUI.

However, when the subsequent driving records of persons whose adjudication was based on three or more DUIs (a subgroup of the one or more DUI group) were compared with those having no DUI, there were significant differences. Table 8 indicates that drivers in the first group had more total subsequent DUIs. In terms of other violations, there were no significant differences between two groups. Although not definitive, these results do not support the early relicensing of habitual offenders whose adjudication is based solely on DUI offenses.

Adjudicated Versus Certified Drivers

Another question pertaining to this DUI population is the effect of the Act on DUI offenders as opposed to non-DUI offenders. To answer this question, the subsequent driving records of adjudicated drivers were compared with those merely certified for three groups: (*a*) drivers whose adjudication or certification was based on non-DUI traffic offenses, (*b*) those whose adjudication or certification was based on at least one DUI, and (*c*) those whose adjudication or certification was based on three or more DUIs.

The impact of the program on habitual offenders with one or more previous DUIs was very similar to the impact on those with no DUIs (see Tables 9 and 10). For both groups, the adjudicated group had fewer subsequent traffic events than did the certified group in each of the 4 years. The adjudicated group also had fewer total subsequent crashes. The certified and adjudicated groups with no previous DUI did not differ in terms of subsequent DUIs. However, the DUI adjudicated group had fewer subsequent DUIs in the fourth year and fewer total DUIs than the DUI certified group. This indicates that the positive impact of the Act was similar for the two groups except that adjudication resulted in fewer subsequent DUIs than certification for the group with one or more previous DUIs.

This was not the case for drivers whose adjudication was based on three or more DUIs. As seen in Table 11, there were no significant differences between the numbers of subsequent DUIs, crashes, traffic events, or time to next traffic event of the adjudicated and certified groups who had three or more DUIs. Thus, although there was a significant positive effect of adjudication shown for drivers with no previous DUI and those with at least one previous DUI, no positive effect was shown for drivers with three or more previous DUIs. This indicates that the Act was not effective in reducing either the amount of driving or the negative consequences of driving for drivers with a serious drinking problem. This finding suggests that persons with one or two DUIs contributing to their adjudication may have benefited from the Act but that the group with three or more DUIs did not.

TABLE 8 Subsequent Driving Records of Adjudicated Drivers With Three or More DUIs (N = 220) and Those Without DUIs (N = 110)[a]

			Deviation from Grand Mean[b]		Significance Testing	
Year	Offense	Grand Mean	No DUIs	Three or More DUIs	F[c]	Significance
First	DUI convictions	.039	–0.03	0.01	1.48	N.S.
	Traffic events	.118	–0.02	0.01	0.30	N.S.
Second	DUI convictions	.021	–0.02	0.01	1.80	N.S.
	Crashes	.006	–0.01	0.01	2.37	N.S.
	Traffic events	.091	–0.07	0.03	2.99	N.S.
Third	DUI convictions	.045	–0.05	0.02	3.70	N.S.
	Crashes	.009	–0.01	0.01	1.45	N.S.
	Traffic events	.100	–0.03	0.02	0.69	N.S.
Fourth	DUI convictions	.033	0.00	0.00	0.00	N.S.
	Crashes	.018	0.00	0.00	0.00	N.S.
	Traffic events	.118	0.00	0.00	0.01	N.S.
Total	DUI convictions	.139	–0.10	0.05	4.77	$p < .05$
	Crashes	.033	–0.03	0.01	1.69	N.S.
	Traffic events	.427	–0.13	0.06	2.02	N.S.
	Time to next event	1,220.876	52.14	–26.07	1.06	N.S

[a]Controlling for previous driving-under-suspension/revocation violations, previous suspensions, and previous failure-to-stop-at-the-scene-of-an-accident (misdemeanor) violations.
[b]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[c]1 degree of freedom.

TABLE 9 Subsequent Driving Records of Adjudicated Drivers (N = 110) and Certified Drivers (N = 156) with No DUIs[a]

Year	Offense	Grand Mean	Deviation from Grand Mean[b]		Significance Testing	
			Certified	Adjudicated	F[c]	Significance
First	DUI convictions	.026	0.00	–0.00	0.26	N.S.
	Traffic events	.259	0.08	–0.11	5.09	$p < .05$
Second	DUI convictions	.034	0.02	–0.02	2.69	N.S.
	Crashes	.019	0.01	–0.02	3.83	$p < .05$
	Traffic events	.248	0.10	–0.15	11.44	$p < .01$
Third	DUI convictions	.019	0.01	–0.01	0.55	N.S.
	Crashes	.041	0.02	–0.03	3.10	N.S.
	Traffic events	.312	0.14	–0.20	15.53	$p < .01$
Fourth	DUI convictions	.049	0.00	–0.01	0.11	N.S.
	Crashes	.038	0.02	–0.03	3.70	N.S.
	Traffic events	.376	0.16	–0.23	13.46	$p < .01$
Total	DUI convictions	.128	0.03	–0.04	2.01	N.S.
	Crashes	.098	0.05	–0.08	8.37	$p < .01$
	Traffic events	1.195	0.48	–0.69	30.62	$p < .01$
	Time to next event	1,023.470	–110.06	156.08	17.12	$p < .01$

[a]Controlling for previous driving-under-suspension/revocation violations and previous suspensions.
[b]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[c]1 degree of freedom.

TABLE 10 Subsequent Driving Records of Adjudicated Drivers (N = 503) and Certified Drivers (N = 506) with One or More DUIs[a]

Year	Offense	Grand Mean	Deviation from Grand Mean[b]		Significance Testing	
			Certified	Adjudicated	F[c]	Significance
First	DUI convictions	.053	0.01	–0.01	2.16	N.S.
	Traffic events	.181	0.04	–0.04	5.62	$p < .05$
Second	DUI convictions	.046	0.01	–0.01	1.46	N.S.
	Crashes	.010	0.00	0.00	0.00	N.S.
	Traffic events	.152	0.04	–0.04	6.04	$p < .05$
Third	DUI convictions	.056	0.01	–0.01	1.42	N.S.
	Crashes	.022	0.01	–0.01	7.76	$p < .01$
	Traffic events	.019	0.08	–0.08	20.26	$p < .01$
Fourth	DUI convictions	.055	0.02	–0.02	7.17	$p < .01$
	Crashes	.035	0.01	–0.01	3.83	$p < .05$
	Traffic events	.184	0.06	–0.07	18.27	$p < .01$
Total	DUI convictions	.209	0.05	–0.05	9.83	$p < .01$
	Crashes	.066	0.02	–0.02	8.96	$p < .01$
	Traffic events	.708	0.22	–0.23	35.93	$p < .01$
	Time to next event	1,141.375	–55.35	55.68	13.42	$p < .01$

[a]Controlling for previous driving-under-suspension/revocation violations, previous suspensions, previous administrative revocations for DUI, previous ASAP attendance, and previous failure-to-stop-at-the-scene-of-an-accident (misdemeanor) violations.
[b]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[c]1 degree of freedom.

DISCUSSION OF RESULTS

Perhaps the most striking finding of this study is that there is a sizable group of individuals who were certified as habitual offenders by DMV in 1986 but who had not been adjudicated as of 1991. Although the existence of this group provided for the methodology used to analyze the effectiveness of the Act on traffic safety, it is somewhat disconcerting that so many individuals are never brought to court on the charge of being a habitual offender. Unfortunately, there was no requirement or means for DMV to track the reasons the certified group had not been adjudicated. However, voluntary reporting by some courts indicated that a substantial portion of the certified group could not be located.

In general, however, the data analyzed in this study indicate that the individuals who are adjudicated under the Act have

TABLE 11 Subsequent Driving Records of Adjudicated Drivers (N = 220) and Certified Drivers (N = 151) with Three or More DUIs[a]

			Deviation from Grand Mean[b]		Significance Testing	
Year	Offense	Grand Mean	Certified	Adjudicated	F[c]	Significance
First	DUI convictions	.049	0.00	0.00	0.04	N.S.
	Traffic events	.129	0.02	–0.02	0.68	N.S.
Second	DUI convictions	.032	0.00	0.00	0.19	N.S.
	Crashes	.011	0.00	0.00	0.05	N.S.
	Traffic events	.113	0.03	–0.02	1.90	N.S.
Third	DUI convictions	.054	–0.02	0.01	1.24	N.S.
	Crashes	.019	0.01	–0.01	2.24	N.S.
	Traffic events	.108	0.01	0.00	0.11	N.S.
Fourth	DUI convictions	.030	0.00	0.00	0.04	N.S.
	Crashes	.032	0.02	–0.01	1.91	N.S.
	Traffic events	.121	0.01	–0.01	0.34	N.S.
Total	DUI convictions	.164	–0.01	0.01	0.18	N.S.
	Crashes	.062	0.03	–0.02	2.65	N.S.
	Traffic events	.472	0.08	–0.05	1.94	N.S.
	Time to next event	1,224.593	–11.50	7.90	0.18	N.S.

[a]Controlling for previous driving-under-suspension/revocation violations, previous suspensions, and previous failure-to-stop-at-the-scene-of-an-accident (misdemeanor) violations.
[b]To obtain the mean for each group, add the deviation from the grand mean to the grand mean.
[c]1 degree of freedom.

more prior DUI convictions than the certified group, even though the certified group have more prior SOL and minor convictions. The data also indicate that adjudication results in the commission of fewer traffic violations and crashes and a longer average conviction-free or crash-free time than only certification. Thus, there is some indication that the Act has a positive impact on traffic safety.

However, three-time DUI offenders fare much worse than others after adjudication and are most likely to have subsequent DUIs. Although Virginia law allows for the early restoration of driving privileges for rehabilitated DUI offenders, it is clear from this analysis that not all DUI habitual offenders deserve special consideration for relicensure. Clearly, guidance should be provided to the courts to assist judges in discriminating between DUI habitual offenders who no longer pose a risk on the highway and those whose driving may still endanger themselves and others.

REFERENCES

1. L. K. Li and P. F. Waller. *Evaluation of the North Carolina Habitual Offender Law.* Highway Safety Research Center, Chapel Hill, N.C., 1976.
2. L. K. Staplin. *Effectiveness of Current Sanctions Against Habitual Offenders.* Ketron, Inc., Malvern, Pa. 1989.
3. C. J. Helander. *An Evaluation of the California Habitual Traffic Offender Law.* Department of Motor Vehicles, Sacramento, Calif., 1986.

The opinions, findings, and conclusions expressed in this paper are those of the authors and not necessarily those of the sponsoring agencies.

Publication of this paper sponsored by Committee on Alcohol, Other Drugs, and Transportation.

Removing the "High" from the Highways: The Impact of Virginia's Efforts To Combat Drug-Related Driving Under the Influence

JACK D. JERNIGAN

Beginning on April 1, 1988, a revision to Virginia law gave police officers the authority to require an individual suspected of drug-related driving under the influence (DUI) to submit a blood sample to be tested for drugs. Concurrent with the implementation of the revised law, Virginia initiated a pilot Drug Recognition Technician (DRT) Program, which concentrates on training police officers to detect the signs of impairment consistent with seven broad categories of drugs. The impact of the revised law and the DRT program on arrests and convictions for drug-impaired driving between 1988 and 1990 was evaluated. In addition, the question of whether there was a spillover effect on alcohol-related arrests and convictions and alcohol-related injury and fatality rates was investigated. Drug-related DUI arrests increased in 1988 but declined somewhat in 1989 and 1990; however, the DUI conviction rate for drug-related cases remained relatively stable. Generally, if a drug was detected, the DUI conviction rate was 40 to 70 percent, depending on the type of drugs detected. If no drug was detected, the DUI conviction rate was less than 25 percent. Although the revised law encouraged officers to make more arrests for drug-related DUI, there is no evidence that it reduced fatalities. Further, even though the DRT program helped increase arrests for drug-related DUI, DRT cases were no more likely than non-DRT cases to result in a conviction. However, there is some evidence that the DRT program had a positive influence on the arrest rate for alcohol-related DUI.

Effective April 1, 1988, Virginia implemented a revised law that prohibits impaired driving. A key provision was that police officers can require an individual suspected of driving under the influence (DUI) to submit a blood sample to be tested for drugs even if an evidentiary breath test for alcohol has been administered. The results of the blood test can be used in court to corroborate an officer's testimony that the suspect had been using drugs and as a supplement to the officer's testimony of the evidence of the suspect's impaired behavior. However, drugs other than alcohol are so chemically complex, and their effects so varied among individuals, that currently there is no scientific way to relate blood drug concentration to blood alcohol concentration (BAC) or to impairment (*1*).

In preparing for the implementation of the revised DUI law, the Virginia Department of Motor Vehicles (DMV) and the Virginia State Police (VSP) established a task force. To supplement the revised law, a pilot Drug Recognition Technician (DRT) Program was established in several jurisdictions, which consisted of training officers to determine whether an individual had used a drug and the class of drug he or she had used. However, since enforcement of the revised statute was not limited to the pilot jurisdictions, a statewide program was developed. The Virginia Division of Forensic Science (DFS) developed and distributed statewide standardized regulations, procedures, forms, and information sheets concerning the submission of blood samples for individuals suspected of driving while impaired by drugs. The revised statute was publicized through a public information campaign. Several policy guidance memoranda were developed and sent to police agencies to encourage enforcement of the revised statute and clarify procedures for its effective use. An additional strategy was to train officers in the use of standardized field sobriety tests. Hence, the strategies mainly comprise an enforcement training program, albeit one supplemented by public information and education efforts.

Virginia Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22903-0817.

LITERATURE REVIEW

Studies of DUI countermeasures involving enforcement have generally found that such efforts can be effective. A number of studies found that enforcement efforts targeting DUI can significantly increase DUI arrests and reduce crash or fatality rates (*2–6*).

Other studies point to the precarious nature of the effectiveness of enforcement programs. Voas and Hause (*7*) found that nighttime crashes, a surrogate measure for alcohol-related crashes, decreased during the implementation of a nighttime DUI special enforcement program. However, the researchers pointed out that the effectiveness of the program was greatest in the early stages of its implementation.

Ross (*8*) investigated the success of the Europeans, particularly the Scandinavians, in deterring drunk driving and concluded that the deterrence effects of these models were not as effective as had been reported. Specifically, although legislative action and other deterrence efforts had an initial impact in reducing drunk driving and fatal crashes, the benefits were only for the short term. Although these countries have stricter laws and harsher penalties than the United States, Ross concluded that social norms are more likely to be at the

root of their success in deterring drunk driving—a conclusion echoed in a later, related monograph by Jacobs (*9*).

Liban et al. (*10*) examined a number of drunk driving countermeasure programs in Canada. They found that a number of community enforcement efforts were attempted but concluded that these efforts had a limited impact on reducing drunk driving. Furthermore, they concluded that the limited effectiveness was short lived.

In the United States, there are some indications that DUI countermeasures may affect fatalility rates, at least in the short run. Hingson et al. (*11*) related the flurry of media and public attention and legislative action focused on DUI in the early 1980s to the drop in fatal crashes between 1980 and 1985. However, they pointed out that this trend soon ended and was, in fact, reversed between 1985 and 1986.

Hingson et al. (*3*) studied the impact of legislation in Maine that made driving with a BAC of 0.10 percent or higher a per se violation of the state's DUI law. They concluded that the legislation did not have a lasting deterrent impact. One reason was that it failed to change drivers' perceptions that their chances of being apprehended and arrested for drunk driving had increased substantially subsequent to the implementation of the law.

PURPOSE AND SCOPE

The primary objective of the study was to determine the effectiveness of Virginia's program to combat drug-impaired driving, particularly the DRT program, in increasing arrests and convictions for drug-impaired driving and decreasing traffic injuries and fatalities. The scope of this evaluation was limited to Virginia's drug-impaired driving program. These data do not address the potential effectiveness of the DRT program as it might be implemented in other states. That is, the DRT program itself is limited by the laws of Virginia, which may differ from the laws of other states.

Arrests examined in this investigation represent only those arrests in which an officer requested and collected a blood sample to be tested for drugs. Because alcohol- and drug-related DUI cases are charged under the same statute in Virginia, there is no way to separate them in the absence of a chemical test.

Another limitation of the study was that in Virginia, as in many states, driving with a BAC of 0.10 percent or higher as shown on an evidentiary breath or blood test is considered per se evidence of impairment. Since alcohol impairment and drug impairment are charged under the same statute, the presence or absence of drugs has little influence on the probability of a DUI conviction in cases with a BAC of 0.10 percent or higher (called high-BAC cases) because a DUI conviction is highly probable given the results of the blood or breath test for alcohol alone. To control for the potential conviction rate bias of considering high-BAC cases in the analysis, the researcher compared only suspected drug-related DUI cases in which either no alcohol was detected or the BAC was less than 0.10 percent (called low-BAC cases). In effect, this method ensured that conviction rates would not be elevated by case selection (i.e., by simply processing a greater number of high-BAC cases through the drug testing laboratory).

METHODOLOGY

If an officer requires a suspect to submit a blood sample, two vials of blood are drawn, normally within 2 hr of the offense, and one is sent to DFS for analysis. Virginia law provides that the second vial may, at the request of the suspect, be sent to an approved laboratory for independent analysis.

The vial that is forwarded to DFS is tested first for alcohol content. If the sample has a BAC level of 0.10 percent or higher, no additional tests are conducted unless a DRT was involved in the arrest. Radioimmunoassay (RIA) is used to screen the blood for evidence of drug use. Gas chromatography/mass spectrometry (GC/MS) is used to confirm all samples that were positive on RIA for any drug. A finding of drugs is reported only for samples that are positive for both RIA and GC/MS. A report of the test results is sent to the local court of jurisdiction.

Using the data collected by DFS, the only central source of information for drug-related DUI arrests, it is possible to track cases back to arrest and forward to resolution. Cases were tracked through at least one of two avenues. First, beginning in summer 1990, court records were checked to ascertain the judicial resolution of each case. Cases that had been resolved and were of record in the local office of the clerk of the court were tracked. Second, in a sample of cases, the arresting officer was contacted and questioned about the resolution of the case.

By cross-tabulations, low-BAC cases were analyzed to determine whether there was a significant relation ($p < .05$) between the year a sample was submitted to DFS and whether a DRT was involved in the case. This method was also used to examine the relation between the laboratory results and whether a DRT was involved in the case.

Next, the researcher examined the DUI conviction rate. DUI convictions included a few that were being appealed when the conviction data were collected. Convictions on non-DUI charges, including a lesser charge of reckless or improper driving, were not counted as DUI convictions but were considered as being resolved. Cross-tabulations were used to determine whether DRT involvement influenced the relation between the DUI conviction rate and (*a*) the year of submission to the DFS and (*b*) the laboratory results.

In addition, the researcher examined whether the emphasis placed on drug-impaired driving had a spillover effect on alcohol-related DUI arrests and convictions and alcohol-related injury and fatality rates. Rates were calculated for each jurisdiction per 1,000 licensed drivers. Using the *t* test, rates for 1986 were compared with the rates for 1990 (2 years before and after the enactment of the revised DUI law) to determine whether DRT and non-DRT jurisdictions differed significantly.

ANALYSIS

Arrests for Drug-Impaired Driving

Table 1 indicates that between April 1, 1988, and December 31, 1990, DFS received 1,199 low-BAC blood samples to be tested for drugs for DUI cases. Overall, 18.3 percent of the samples submitted to DFS between 1988 and 1990 involved

TABLE 1 Number of Drug-Related DUI Cases by Year Submitted to DFS: BAC < 0.10 percent*

TYPE OF CASE	1988	1989	1990	TOTAL
No DRT Involved	286 (74.3%)	350 (84.7%)	343 (85.5%)	979 (81.7%)
DRT Involved	99 (25.7%)	63 (15.3%)	58 (14.5%)	220 (18.3%)
TOTAL	385	413	401	1,199

*Significant at $p < .05$.

a DRT. In 1988, the first year in which the revised law was in effect, DRTs were involved in 25.7 percent of all samples received by DFS; by 1990, this percentage had dropped to 14.5. This decline was statistically significant. In addition, even though the revised DUI law was in effect for only 9 months in 1988, there were more low-BAC DRT submissions in that year than in either of the 2 subsequent years.

Table 2 indicates that there is a significant relation between whether a DRT was involved in a case and the type of drug that was detected. In particular, cases in which PCP was detected were vastly more likely to be non-DRT cases. Non-DRT cases were more likely to involve multiple drugs. Overall, drugs were detected in 64.6 percent of the samples; however, "no drugs detected" does not necessarily mean that there was no drug present. It is possible that a drug was present for which no test was available, a drug was present but at a concentration too low to be confirmed by DFS (e.g., the dosage level of LSD is too low to be confirmed), or a drug was present at the time of the traffic stop but had metabolized or dissipated before the blood sample was taken.

Convictions for Drug-Impaired Driving

As seen in Table 3, there is a significant relation between the year the sample was submitted to DFS and the result of the case. The percentage of cases resulting in a DUI conviction was highest in 1988, declined in 1989, and declined further in 1990. Table 3 further indicates that the DUI conviction rate of DRT cases remained relatively stable, around 40 percent, but that the DUI conviction rate for non-DRT cases decreased from more than 50 percent in 1988 to less than 37 percent in 1990.

Table 4 indicates that there is a significant relation between the laboratory results and whether a case resulted in a conviction. If a drug was detected in the sample, the overall DUI conviction rate ranged from about 40 to 70 percent, depending on the drug detected. However, if no drug was detected and alcohol was detected at a level less than 0.10 percent BAC, there was less than a 25 percent DUI conviction rate. Finally, when neither drugs nor alcohol was found, less than 15 percent of the cases resulted in a DUI conviction.

Alcohol-Impaired Driving

Table 5 indicates that the alcohol-related DUI arrest rate for 1,000 licensed drivers overall and among the non-DRT jurisdictions declined significantly from 1986 to 1990. However, in the DRT jurisdictions, there was no significant difference in the arrest rates of 1986 and 1990. Table 5 also indicates that the average conviction rate for DRT jurisdictions in-

TABLE 2 Number of Drug-Related DUI Cases by Laboratory Result: BAC < 0.10 percent*

LABORATORY RESULT	NO DRT INVOLVED	DRT INVOLVED	TOTAL
Multiple drugs	193 (19.7%)	23 (10.5%)	216 (18.0%)
Marijuana	150 (15.3%)	48 (21.8%)	198 (16.5%)
PCP	160 (16.3%)	7 (3.2%)	167 (13.9%)
Cocaine	71 (7.3%)	16 (7.3%)	87 (7.3%)
Other drugs	79 (8.1%)	28 (12.7%)	107 (8.9%)
No drugs detected, low BAC	55 (15.8%)	50 (22.7%)	205 (17.1%)
No drugs detected, no BAC	171 (17.5%)	48 (21.8%)	219 (18.3%)
TOTAL	979	220	1,199

*Significant at $p < .05$.

TABLE 3 Drug-Related DUI Conviction Rate by Year Submitted: BAC < 0.10 percent

YEAR	NO DRT INVOLVED*	DRT INVOLVED	TOTAL*
1988	50.7% (n = 201)	40.5% (n = 84)	47.7% (n = 285)
1989	44.2% (n = 260)	38.0% (n = 50)	43.2% (n = 310)
1990	36.8% (n = 253)	42.4% (n = 33)	37.4% (n = 286)
TOTAL	43.4% (n = 714)	40.1% (n = 167)	42.8% (n = 881)

*Significant at $p < .05$.

TABLE 4 Drug-Related DUI Conviction Rate by Laboratory Result: BAC < 0.10 percent

LABORATORY RESULT	NO DRT INVOLVED*	DRT INVOLVED*	TOTAL*
Multiple drugs	61.3% (n = 142)	52.9% (n = 17)	60.4% (n = 159)
Marijuana	46.1% (n = 115)	71.4% (n = 35)	52.0% (n = 150)
PCP	69.3% (n = 114)	40.0% (n = 5)	68.1% (n = 119)
Cocaine	42.0% (n = 50)	71.4% (n = 14)	48.4% (n = 64)
Other drug	42.6% (n = 61)	30.0% (n = 20)	39.5% (n = 81)
No drugs detected, low BAC	23.4% (n = 111)	25.0% (n = 40)	23.8% (n = 151)
No drugs detected, no BAC	14.9% (n = 121)	13.9% (n = 36)	14.6% (n = 157)
TOTAL	43.4% (n = 714)	40.1% (n = 167)	42.8% (n = 881)

*Significant at $p < .05$.

TABLE 5 Average DUI Arrests and Convictions per 1,000 Licensed Drivers

Rate	1986	1990	% Change
Average DUI Arrest Rate			
Non-DRT Jurisdictions*	14.19	12.41	–12.5
DRT Jurisdictions	13.41	13.48	+ 0.5
All Jurisdictions*	14.16	12.46	–12.0
Average DUI Conviction Rate			
Non-DRT Jurisdictions	10.84	10.42	– 3.9
DRT Jurisdictions	11.45	12.19	+ 6.5
All Jurisdictions	10.87	10.50	– 3.4

*Significant at $p < .05$.

creased between 1986 and 1990, although the rates for non-DRT jurisdictions declined slightly. However, these changes in conviction rate were not statistically significant.

Table 6 indicates that the injury rate for alcohol-related crashes decreased significantly between 1986 and 1990 in both DRT and non-DRT jurisdictions. On the other hand, there was no significant change in the alcohol-related fatality rate between 1986 and 1990.

DISCUSSION OF RESULTS

Arrests for Drug-Impaired Driving

The number of blood samples to be tested for drugs for low-BAC cases declined from an average of more than 42 per month in 1988 to an average of fewer than 35 per month in 1989 and 1990. Thus, the number of cases that might have been pursued as a consequence of the revised law declined in the second and third years of the law's implementation, as did the average number of cases submitted by DRTs. These findings are consistent with the literature on enforcement programs. In particular, many enforcement efforts begin by moving toward accomplishing their goals, but the initial emphasis as well as the initial success begins to diminish.

In 1988, two chiefs of police were actively involved in developing the DRT program and in working with the task force; by 1990, no chief of police was actively involved. Instead, the program and its development had been allocated to lower administrative levels of the enforcement agencies involved.

In addition, there were initially two sergeants who were among the first to receive DRT training and were the leaders and chief salespeople for the DRT program. For differing reasons, both moved from their initial responsibilities in overseeing the program to other duties. Without these sergeants and the chiefs, the program lost much of its continuity and leadership.

At DMV, responsibility for this program moved from a level of involvement by relatively high management to a lower level of training coordinators. Hence, contact with police agencies came from the lower levels of the agency. Likewise, early in the program's development, there was a flurry of public information and education activity that had all but ceased by 1989.

Statewide, the task force sent out several policy guidance memoranda that were intended to inform officers about the revised law and provide suggestions for pursuing cases under the revised law. The last of five memoranda was sent out on March 7, 1989. The task force also held bimonthly or quarterly meetings in 1988 and 1989, but few have been held since.

No individual or agency is necessarily to blame for the drop in the number of submissions to DFS, the decline in activity, or the delegation of authority. Rather, the drop is characteristic of a program of this type running through a life cycle of enthusiasm to decline. As much of the literature points out, any success of an enforcement program is usually short lived. Much of the success of Virginia's DRT program seems to have likewise been short lived.

Convictions for Drug-Impaired Driving

Between 1988 and 1990, the conviction rate for non-DRT cases declined from more than 50 to less than 37 percent. Initially, conviction rates were higher for non-DRT cases than for DRT cases. On the other hand, DRT cases had a relatively stable conviction rate of about 40 percent throughout the 3 years. During the first 2 years, the differences between the conviction rates for DRTs and non-DRTs could largely be explained by the fact that most PCP cases were non-DRT cases (*12*). That is, a PCP case was more likely than any other case to result in a conviction. Hence, the existence of a substantial number of PCP cases in the non-DRT sample inflated the conviction rate for non-DRTs. When the DRT and non-DRT samples were made more comparable by consideration of only non-PCP cases, the difference in conviction rates was eliminated (*12*). Similarly, a drop in the number of PCP cases submitted by non-DRTs in 1989 and 1990 likely functioned to decrease the overall conviction rate for non-DRT cases simply because any other laboratory result was associated with a lower conviction rate than a finding of PCP.

Finally, although different laboratory results were related to different conviction rates, a laboratory result of "no drugs detected" was associated with a conviction rate of less than 25 percent.

Alcohol-Impaired Driving

Between 1986 and 1990, the alcohol-related DUI arrest rate per 1,000 licensed drivers for non-DRT jurisdictions declined significantly, but that for DRT jurisdictions remained stable. This indicates that there may have been a spillover effect of the DRT program on alcohol-related DUI arrests. That is,

TABLE 6 Average Alcohol-Related Injuries and Fatalities per 1,000 Licensed Drivers

Rate	1986	1990	% Change
Average Alcohol-Related Injury Rate			
Non-DRT Jurisdictions*	4.22	3.65	–13.5
DRT Jurisdictions*	4.59	3.49	–24.0
All Jurisdictions*	4.23	3.65	–13.7
Average Alcohol-Related Fatality Rate			
Non-DRT Jurisdictions	0.20	0.19	– 5.0
DRT Jurisdictions	0.10	0.11	+10.0
All Jurisdictions	0.19	0.19	—

*Significant at $p < .05$.

by concentating some training and enforcement on drug-impaired driving, it is possible that the DRT jurisdictions helped fight off a decline in the DUI arrest rate in the non-DRT jurisdictions.

Although the DUI conviction rate per 1,000 licensed drivers for non-DRT jurisdictions declined and the rate for DRT jurisdictions increased, neither change was significant. Thus, there is no indication that either the revised law or the DRT program affected the DUI conviction rate as measured per 1,000 licensed drivers.

The alcohol-related injury rate per 1,000 licensed drivers was down in both DRT and non-DRT jurisdictions, but the decline was greater in DRT jurisdictions. Thus, the existence of the DRT program and the stable arrest rates for DUI may have decreased the alcohol-related injury rate in DRT jurisdictions. However, because there was not a significant change in the fatality rate, there is no evidence that the revised law or the DRT program had any impact on reducing traffic fatalities.

CONCLUSIONS

Between 1973 and 1984, there was an average of only 11 convictions per year for drug-related DUI in Virginia (*13*). Thus, the revised law was effective in increasing the absolute number of arrests and convictions for drug-related DUI. Furthermore, because DRTs make up less than 1 percent of the statewide enforcement strength and were involved in about 15 percent of the drug-related DUI cases, there is evidence that DRT training increased the level of law enforcement. In addition, because the alcohol-related DUI arrest rate remained stable in DRT jurisdictions and declined in other jurisdictions, there is evidence that the DRT program may have had a spillover effect on the enforcement of alcohol-related DUI.

However, there are some issues of concern about Virginia's efforts to combat drug-related DUI. The overall drug-related DUI conviction rate in both DRT and non-DRT jurisdictions is only about 40 percent, although the conviction rates for cases in which a particular drug (e.g., cocaine, marijuana, and PCP) was detected are higher. Thus, if the conviction rate is to increase, efforts are needed to strengthen many DRT and non-DRT drug-related DUI cases. There is no evidence that either the 1988 law or the DRT program functioned to decrease the fatality rate. Most disconcerting is the evidence of decline in even the positive measures of short-term success. Drug-related DUI arrests declined after 1988, and emphasis on enforcement and task force and public information and education activities has also diminished. It was concluded that, unless substantially revitalized, Virginia's efforts to combat drug-related DUI will continue to follow the path of decline that has plagued so many other enforcement programs.

REFERENCES

1. J. Mörland. Psychoactive Drugs and Driving Performance. *Proc., 35th International Congress on Alcoholism and Drug Dependence*, Vol. 3, National Directorate for the Prevention of Alcohol and Drug Problems, Oslo, Norway, 1989, pp. 401–409.
2. D. Foley. Case Study in DWI Countermeasures. In *Stop DWI: Successful Community Responses to Drunk Driving* (D. Foley, ed.), Lexington, Lexington, Mass., 1986.
3. R. Hingson, T. Heeren, D. Kovenoch, T. Mangione, A. Meyers, S. Morelock, R. Lederman, and N. Scotch. Effects of Maine's and Massachusetts' 1982 Driving Under the Influence Legislation. *American Journal of Public Health*, Vol. 77, 1987, pp. 593–597.
4. J. Lacey, L. Steward, L. Marchette, P. Popkin, R. Murphy, R. Luche, and R. Jones. *Enforcement and Public Information Strategies for DWI General Deterrence: Arrest Drunk Driving—The Clearwater and Largo, Florida Experience.* DOT Report HS 807 066. U.S. Department of Transportation, 1986.
5. G. Sykes. Saturated Enforcement: The Efficacy of Deterrency and Drunk Driving. *Journal of Criminal Justice*, Vol. 12, 1984, pp. 185–197.
6. R. Voas, A. Rhodinizer, and C. Lynn. *Evaluation of Charlottesville Checkpoint Operations.* DOT Report HS 806 989. U.S. Department of Transportation, 1986.
7. R. Voas and J. Hause. Deterring the Drinking Driver of the Stockton Experience. *Accident Analysis and Prevention*, Vol. 19, 1987, pp. 81–90.
8. H. Ross. *Deterring the Drinking Driver.* Lexington, Lexington, Mass., 1982.
9. J. Jacobs. *Drunk Driving: An American Dilemma.* University of Chicago Press, Chicago, Ill., 1989.
10. C. Liban, E. Vingilis, and H. Blefgen. The Canadian Drinking-Driving Countermeasure Experience. *Accident Analysis and Prevention*, Vol. 19, 1987, pp. 159–181.
11. R. Hingson, J. Howland, S. Morelock, and T. Heeren. Legal Interventions To Reduce Drunken Driving and Related Fatalities Among Youthful Drivers. *Alcohol, Drugs and Driving*, Vol. 4, 1988, pp. 87–98.
12. J. Jernigan. *Virginia's Program To Combat Drug-Related DUI: 1988–1989.* VTRC Report 92-R9. Virginia Transportation Research Council, Charlottesville, 1992.
13. E. Paltell and M. Booz. *Combating the Drug-Impaired Driver: A Prescription for Safer Highways.* VTRC Report 86-R20. Virginia Transportation Research Council, Charlottesville, 1985.

The opinions, findings, and conclusions expressed in this paper are those of the author and not necessarily those of the sponsoring agencies.

Publication of this paper sponsored by Committee on Alcohol, Other Drugs, and Transportation.

Testing the Social-Stress Prevention Model in an Inner-City Day Camp

GARY L. WINN, D. FLOYD JONES, AND CURTIS JAY BONK

In an effort to improve inner-city children's knowledge of problems of drug and alcohol abuse, especially related to traffic safety, a test of the Rhodes and Jason social-stress prevention model was conducted as part of a comprehensive summer day camp for largely minority youngsters in Wheeling, West Virginia. As part of a multidisciplinary and comprehensive approach to youth development called Project YES, the drug and alcohol module (DAM) was designed to improve students' awareness of issues, consequences, and physiological effects. Important findings are that intense drug and alcohol awareness training could be successfully integrated into an intense day camp context among inner-city youth and that the DAM was a reasonably popular part of the curriculum. An 18-item, matched-subject knowledge test showed significant knowledge gains as a result of the intervention. Further exploration into knowledge-residual and actual behavior change is suggested.

Abuse of alcohol and drugs in the United States has captured the attention of parents, teachers, community leaders, and policy makers as never before. Epidemiological studies continue to demonstrate the incredible consequences of drug and alcohol abuse, and research summaries suggest that youths are especially at risk. At least 56 percent of the nation's school-aged children begin to experiment with alcohol before entering high school; by their senior year, the proportion has grown to 92 percent, and approximately one-third of those people were reported to be regular users (*1*). The same authors state, "When this [alcohol] abuse is combined with tobacco, drug drug-taking behavior appears to be well entrenched in our public schools [even though] recent studies indicate a downward trend in drug use in schools."

The National Safety Council (*2*) reports that 1,100 pedestrians under 4 years of age were fatally injured in 1990; 650 fatal pedestrian injuries occurred to children 5 to 14 years old. Children riding bicycles accounted for 400 fatal injuries to persons under 4 years old; 350 fatal bike injuries were tallied for riders 5 to 14 years of age (*2*). The pervasiveness of alcohol and drugs in schools suggests that some of these traffic deaths involved impaired riders or pedestrians.

Social influences to drink or take drugs, especially influences associated with peer groups, suggest that the effect of friends may be so pervasive as to cancel out the threat of legal consequences or parental guidance (*3,4*). Jessor found that youth are more influenced by friends than parents (*5,6*). In a similar context, McPherson et al. (*7*) found that young motorcyclists are significantly more likely to attend a safety course if influenced by peers than if influenced by parents or any form of media. Youth are often unable to even identify the amount of alcohol likely to affect them physiologically (*7,8*) or levels that impair driving performance (*9*).

Inner-city residents may have a greater problem than suburban residents. In a study of emergency room visits, black males had a higher percentage of emergency room episodes caused by drug use than white males. Blacks experienced more emergency room episodes than Hispanics and whites, according to a study of 27 metropolitan areas by the National Institute on Drug Abuse (*10,11*).

Drug use as a social-stress coping mechanism retards the development of further social competencies; failures in adaptive behaviors are related to heavy drug use, according to Wrubel et al. (*12*). Deprived social settings, by further implication, suggest increased drug and alcohol use, and this has been borne out by empirical study. In 1979, Padilla studied black Americans and found higher drug use. Similarly, Bobo found increased drug use among native Americans (*13*).

MODELS OF PREVENTION

The earliest alcohol/drug abuse programs were implemented exclusively in schools and did not focus on traffic safety as a separate sphere of concern. These information (cognitive) models were basically educational curricula emphasizing increased individual responsibility: preparing young people with information and using scare tactics (*14*). Some authors suggest that single-modality programs have limited impact on preventing drug and alcohol abuse (*15–17*).

A more recent set of approaches has evolved during the late 1970s, especially the affective models, positing that drug and alcohol problems result from lack of social competence and clear social values (*18*).

Environmental programs, the most recent set of models proposed in the literature, have an information component and often stress social competence, but they add cultural and even community support to the program and may focus on specific spheres of concern, such as traffic safety problems of incipient drivers. This "broader view" of abuse problems suggests that "appreciation for the complexity of the [abuse] problem holds significant promise" (*14*).

Rhodes and Jason present a compelling model of prevention adaptable for use with inner-city, socioeconomically deprived youth (*19*). Their 1987 social-stress model emphasizes ecological, or environmental, aspects of adolescent behavior. These aspects include three categories of influence: the social

G. L. Winn, Department of Safety Studies; D. Floyd Jones, Department of Sport and Exercise Studies; and C. J. Bonk, Department of Educational Psychology; West Virginia University, Morgantown, W.Va. 26506.

network, social competence, and social resources. These three categories of stress moderators yield a subjective alcohol or drug risk value for a given individual.

In 1988, Rhodes and Jason suggested that stress may arise from within the family, the school, a peer group, or the community. To generate these positive social networks, a child needs adult model figures and must incorporate at least some of their standards and values. Likewise, the influence of the peer group and family is important in reducing social stresses, thereby leading to "a sense of efficacy and control in his or her interactions" (*20*).

Similarly, these authors advise that social competencies, such as decision-making skills and peer pressure resistance, and, by implication, self-esteem, are avenues toward improved self-control. The third critical component in the social-stress model is the social resource factor, which takes the form of repositories of information. The school, the neighborhood, the church, or the community learning center can all serve as social resources for factual information, and even as role models.

Rhodes and Jason conclude, "The social-stress model provides a needed basis for a more comprehensive, ecological approach to prevention. Among the multiple factors that can potentially influence adolescents' alcohol and other drug problems, comprehensive, community-based prevention programs are in our view, the next logical step" (*20*).

THE PROJECT YES MODEL OF PREVENTION

Based on a successful pilot program tested in Charleston, West Virginia, during summer 1990, Project YES (Youth Enrichment Services) was contracted in mid-1991 as a comprehensive, community-based program concentrating a great deal of its focus on drug and alcohol education and traffic safety. Project YES provided the test-bed for the Rhodes and Jason model of prevention.

The three focal components of the 6-week Project YES were (*a*) improving life skills, such as decision making, and the self-esteem of inner-city children 7 to 14 years old; (*b*) creating competencies in literacy, communicating, and writing through personal computers; and (*c*) improving safety and health skills and drug and alcohol knowledges.

Project YES was molded to correspond closely to the three stress moderators closely consistent with those proposed by Rhodes and Jason.

For social networks, Project YES provided both community leaders and teenage mentors as instructors and trained them for 10 hr in programs and necessary skills. Parents were used as adjunct or assistant instructors (a method that proved particularly useful for maintaining discipline during the hot, outdoor day camp). Multicultural instructors from local colleges and institutions provided stories, physical education, and dance instruction. Children learned who could provide factual answers to questions about drugs and alcohol, bicycle safety, computers, and so on.

For social competencies, Project YES provided experts from local high schools who were themselves further trained in self-esteem improvement and decision skill enhancement and also in the inner workings of Project YES. Children learned better interpersonal communication via specific, one-on-one instruction and also by computers (small work groups created their own Project YES newsletters). Bonding to group norms was encouraged through organized group activities (parachute movement, dances, "winning" mock debates, the use of Project YES T-shirts, and so on).

For social resources, Project YES incorporated the use of every available community resource on alcohol/drug information, bicycle safety, computer information, personal hygiene information, academic information, and much more. This in turn was delivered by African-American police officers (motorcycle and cruiser), storytellers from the local library, a farm couple who allowed horseback riding only with appropriate safety gear, church leaders, and parents who originally developed the idea of a community learning center where the Project YES day camp was held. Resource personnel were volunteers and were carefully instructed beforehand about Project YES and why their own position in the community was important to long-term behavior change long after the day camp had ended. Among the community dignitaries attending one or more Project YES activities were Gaston Caperton, governor of West Virginia, and his wife, Rachael Worby.

Figure 1 shows how the parent model of Rhodes and Jason was equipped for testing by the components of Project YES.

RESEARCH METHODS: A MODEL-BASED DRUG AND ALCOHOL CURRICULUM

As a drug and alcohol awareness intervention, the drug and alcohol module (DAM) included specific instruction, arranged in five lessons, as follows:

1. Pre-test, differences between drug use and drug abuse, problems encountered by well-known sports figures, definitions of drugs, types of drugs and alcohol and where they are found, and problems of drinking and driving (2 days);
2. Physiological changes with drug and alcohol use, neighborhood sports figures with drug problems, and common drugs and forms of alcohol (hands-on, local resources) (2 days);

Moderator Type *(Rhodes and Jason)*	**Moderator Component** *(Project YES)*
Social Network	• adult role models • teen mentors • parent figures • consistent adult support • networking outside classroom
Social Competencies	• decision-making training • communications instruction • peer-pressure resistance training • self-esteem instruction
Social Resources	• non-traditional teaching venues (zoo, farm, college, library, public transportation, restaurants)

FIGURE 1 Project YES stress-moderator categories.

3. Social and economic consequences of drug abuse (divorce, jail, problems in school, lost driver's license), child/spouse abuse, and introduction to role-playing (2 days);
4. Special problems (drunk driving and consequences of being caught while driving impaired) and role-playing strategies (1 day); and
5. Review: differences in alcohol sources (liquor, beer, wine), effects of drugs and alcohol on inexperienced people, "gateway" drugs, guest (black police officer), continue role-playing, post-test (3 days).

The DAM relied heavily on role-playing activities (which often included the teen mentors who lived in the local community), strategies to reduce peer pressure (tell an adult, walk away, "just say no") and responding with facts during a "confrontation." To reinforce the resource component of the model, the DAM also included training about where in the Wheeling community to get help against peer pressure and where to get the facts about alcohol and drugs in the community.

The DAM emphasized national and local sports figures who had lost their careers, or worse, through drug or alcohol abuse. The positive impact among Project YES youth of sports prodigies who avoided contact with drugs and alcohol could not be overstated. These included Jim Paige, East Wheeling's own college basketball star (now West Virginia's state treasurer).

As much as possible, the concepts from the DAM were referenced during other Project YES components such as self-esteem training, computer-based writing activities, and nutrition projects. For example, students might write about a sports figure affected by drugs or compose a story focusing on a peer-pressure tactic during computer workshops.

An 18-item knowledge pre-test was developed and administered to 82 children 6 to 14 years old enrolled in the 1991 Project YES day camp in East Wheeling, West Virginia. The pre-tests were administered in late June on 2 consecutive days, followed by the 3-week instructional period. The same test was administered as a post-test on 2 consecutive days in mid-July 1991. Thirty-one of the students who took the pre-test were in attendance for the post-test. Participation in all Project YES activities was voluntary; thus, the DAM instructional group changed slightly each day over the instructional period. For this reason, a matched-subject design was chosen to reduce the need for a true experimental control group and to account for what appear to be dropouts. The voluntary nature of the program made a true control group logistically impossible.

RESULTS: EFFECTIVENESS MEASURES

Knowledge about drugs and alcohol showed significant improvements in the 3-week period of the DAM for all participants. On an 18-question inventory, 31 students averaged 7.90 questions correct on the pre-test but 13.81 items correct on the post-test. The improvement in pre- to post-test scores was statistically significant at $p < .001$, suggesting that the application of the DAM was responsible for the improvement in post-test scores. All students except one had improved post-test scores; the single exception had scores that did not change.

Table 1 gives descriptive pre- and post-test breakdowns of score means within the variables age and gender on the DAM. (The "young" group consists of ages 6 to 11 years; "old" indicates 12 to 14).

The next analysis concerned the gain scores between pre- and post-tests. The anticipated significant gains were observed (see Table 2).

The following analysis revealed that whereas older students gained slightly more knowledge than younger students, there was no real age difference in drug and alcohol score gains.

However, the results indicated that on entering the DAM, there was a large raw score performance gap related to age. The older group performed significantly better on the pre-test than the younger group ($p < .0015$). The gap still existed on the post-test; results of the post-test only indicated that young and old groups performed significantly differently ($p < .0003$) (see Table 3).

The young group of 11 students improved knowledge scores significantly ($p < .0001$) and the older group of 20 students also performed better on the post-test ($p < .0001$). Improvements were observed, therefore, regardless of student age.

A final analysis indicated that the subtracted differences between pre- and post-tests were slightly higher for females, but not significantly higher than the males. The female-only group showed significant improvement on the post-test ($p < .0001$), and this feature was also observed for the male group ($p < .0001$).

TABLE 1 Descriptive Raw Score Breakdown, Drug and Alcohol Module

Gender	Score	Std. Dev.	Age	Score	Std. Dev.
Pre-test			Pre-test		
Female	8.85	3.88	Young	4.90	1.92
Male	7.11	4.36	Old	9.54	4.12
Post-test			Post-test		
Male	14.92	5.26	Young	9.90	3.50
Female	12.89	4.48	Old	15.95	4.10

TABLE 2 Summary, Hypothesis Test of Difference Scores

Sample Size	31
Sample Mean	5.98
Standard Dev.	3.82
Standard Error	.541
T value	-10.98
df	38
2-tailed prob.	<.0001

The DAM was rated the fourth-most-popular activity in Project YES, behind field trips, bicycle safety activity, and computer instruction and ahead of self-esteem, hygiene, dance, and nutrition.

DISCUSSION OF RESULTS

Our results suggest that when applied in a context of a community-based, comprehensive youth development program, a 3-week drug and alcohol awareness module can foster significant improvements in knowledge of problems of abuse, legal consequences of drinking and driving, and physiological responses. We note that all students improved (one student held the same score) from preintervention test to postintervention test, and the analysis strongly implies that the intervention was responsible for the gains.

However, because a true control group was impossible, the possibility remains that the students who elected not to return to Project YES during times when the post-test was given were less motivated. This alternative explanation cannot be countered.

That older students performed significantly better than younger students on both the pre- and post-tests may not be surprising, since the issues are more germane to older children, who have greater cognitive skills available anyway. Age differences were observed on both pre- and post-tests, and older students were better performers all around. The DAM is age specific but not gender specific, since males and females performed approximately the same on pre- and post-tests.

TABLE 3 Summary, Hypothesis Test for Age, Pre- and Post-Tests

	Pre-test		Post-test	
	Young	Old	Young	Old
Mean	4.98	9.55	9.98	12.88
Std. Dev	1.92	4.12	3.59	4.19
t value	-3.58		-4.06	
df	29		29	
2-tailed prob.	<.0015		<.0003	

From the point of view of the social-stress model, it appears that students between 6 and 14 years old of both sexes can significantly improve their awareness of drug and alcohol problems in a traffic safety context as a result of a structured, nurturing social system complete with the three necessary components: social networks, social competencies, and social resources.

CONCLUSIONS AND RECOMMENDATIONS

The investigators make no claim that these results are the "final answer" to solving drug and alcohol problems among youth who are also bicycle riders and future drivers. However, these results indicate that formal safety and health training is, indeed, possible in the inner city and that, under the right conditions, it is a popular part of an integrated approach to the problem. At a minimum, these results warrant further research based on the prevention model used here and among inner-city youth.

The investigators are confident that the results of administering the DAM within the context of a larger, supportive youth development program such as Project YES has successfully improved awareness of certain components of the drug and alcohol problem. If improving long-term behavior is the ultimate goal, a strong foundation of knowledge can support changes in attitude and skills that show up later.

However, even though the results of this test of the Rhodes and Jason model do not support a conclusion that the observed effectiveness of the DAM in improving awareness necessarily leads to improved behaviors (declining to drive a car after drinking, for example) or skills (in peer-pressure defense, for example), it is entirely logical to encourage further tests of the model on behavior and skill change over the long term.

Because of the ineffectiveness of unimodal programs seeking to change drug and alcohol awareness levels through single, personal characteristics (self-esteem or peer-pressure training, but not both), it seems that the utility of the Rhodes and Jason model (or similar multimodal, integrated approach) is bolstered.

It could be argued that improvements in a single sphere of Project YES will only be effective in the presence of the others; in fact, this is what Project YES planners had hoped for. The comprehensive nature of Project YES meant that each facet of youth development was linked, reinforced, tied, and otherwise valued by each other module within the community. This is perhaps the reason for the observed success of the DAM: when explicitly supported by self-esteem training, role modeling, parental involvement, peer-pressure training, and decision making, improvements in drug and alcohol awareness appear. Unlike other models that might anticipate improvements in alcohol and drug awareness as a function of three independent dimensions of contact, the Rhodes and Jason model is far more predictive and useful.

REFERENCES

1. L. D. Johnston, P. M. O'Malley, and J. G. Bachman. *Drug Use, Drinking and Smoking: National Survey Results from High School,*

College, and Young Adult Populations. National Institute on Drug Abuse, Rockville, Md., 1975–1988.
2. *Accident Facts* (1991 edition). National Safety Council, Chicago, Ill., 1991.
3. B. Finley. Drinking and Driving Behaviors in Grades 1–10 in Rural High School. *Current Issues in Alcohol and Drug Nursing: Research, Education and Clinical Practice,* 1983.
4. H. Blane. Problem Drinking in Delinquent and Non-Delinquent Adolescent Males. *The American Journal of Drugs and Alcohol,* Vol. 41, No. 3, 1988, pp. 221–232.
5. R. Jessor and S. Jessor. *Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth.* Academic Press, New York, 1977.
6. R. Jessor. Risky Driving and Adolescent Problem Behavior: An Extension of Problem Behavior Theory. *Alcohol, Drugs and Driving,* Vol. 3, 1987.
7. K. McPherson, G. L. Winn, B. Bailey, et al. *Motorcycle Safety Education Marketing Evaluation.* U.S. Department of Transportation, 1990.
8. M. Forney, P. Forney, H. Davis, et al. Factors Affecting Knowledge, Attitude and Behavior Toward the Use of Alcohol Among Middle School Students. *Proc., American Educational Research Association Annual Conference,* 1984.
9. R. Hetherington, J. Dickinson, K. Cipywnyk, et al. Attitudes and Knowledge About Alcohol Among Saskatchewan Adolescents. *Canadian Journal of Public Health,* Vol. 70, No. 4, 1979, pp. 247–259.
10. D. N. Nurco, J. C. Ball, Shaffer, et al. The Criminality of Narcotic Addicts. *The Journal of Nerve and Mental Disorders,* Vol. 173, 1985, pp. 94–102.
11. J. A. Inciardi. Crack Cocaine in Miami. Presented at NIDA Technical Review Meeting on the Epidemiology of Cocaine Use and Abuse. National Institute on Drug Abuse, Rockville, Md., 1988.
12. J. Wrubel, P. Brennan, and R. S. Lazarus. Perspectives of Stress Coping. In *Social Competence* (J. D. Wine and M. D. Smye, eds.), Guilford Press, New York, 1981.
13. J. K. Bobo. Preventing Drug Abuse Among American Indian Adolescents. In *Preventing Social and Health Problems Through Life Skills Training* (L. D. Gilchrist and S. P. Schinke, eds.), University of Washington Press, Seattle, 1986.
14. L. Wallack and K. Corbett. Illicit Drug, Tobacco and Alcohol Use Among Youth: Trends and Promising Approaches in Prevention. In *Youth and Drugs: Society's Mixed Messages.* Office of Substance Abuse Prevention Monograph 6. U.S. Department of Health and Human Services, 1990.
15. J. D. Hawkins, D. M. Lishner, R. F. Catalano, et al. Childhood Predictors of Adolescent Substance Abuse: Toward an Empirically Grounded Theory. In *Childhood and Chemical Abuse,* Haworth Press, Binghampton, N.Y., 1986, pp. 11–48.
16. J. M. Moskowitz. The Primary Prevention of Alcohol Problems: A Critical Review of the Research Literature. *Journal of Studies on Alcohol,* Vol. 50, No. 1, 1989, pp. 54–88.
17. N. S. Tobler. Meta-Analysis of 143 Adolescent Drug Prevention Programs: Quantitative Outcome Results of Program Participants Compared to Control Comparison Group. *Journal of Drug Issues,* Vol. 16, No. 4, 1986, pp. 537–567.
18. R. Dembo. Key Issues and Paradigms in Drug Use Research. *Journal of Drug Issues,* Vol. 16, No. 1, 1986, pp. 1–4.
19. J. E. Rhodes and L. A. Jason. *Preventing Substance Abuse Among Children and Adolescents.* Pergammon Press, New York, 1988.
20. J. E. Rhodes and L. A. Jason. The Social-Stress Model of Alcohol and Other Drug Abuse: A Basis for Comprehensive, Community-Based Prevention. In *Prevention Research Findings,* Monograph 3, Office of Substance Abuse Prevention, U.S. Department of Health and Human Services, 1987.

Publication of this paper sponsored by Committee on Alcohol, Other Drugs, and Transportation.

Tort Liability Related to Utility Pole Accidents in Florida

FAZIL T. NAJAFI, FADI EMIL NASSAR, AND PAUL KACZOROWSKI

Although many statistical data are gathered on pole accidents, very little information is available on the tort liability associated with utility pole accidents. Tort information is difficult to obtain. No organizations keep track of tort claims against public agencies, and utilities are generally reluctant to disclose tort information for fear of jeopardizing their defense against similar claims. A continuing study conducted at the University of Florida to assess pole-related tort liability in Florida is reported. It was found that for the state and most cities, tort liability related to pole accidents is minimal. Utilities' liability presents a mixed picture. Florida court rulings related to pole accident claims have so far been favorable to utilities and public agencies. If this situation changes, however, liability cost can increase significantly in the future, as is the trend in some other states. Therefore, a prudent policy to deal with pole-related tort liability consists of developing guidelines to relocate hazardous poles and to maintain a data base containing tort information and summaries of related court rulings.

State transportation agencies have to deal with a large number of tort claims every year. This coincides with increasing emphasis on highway safety by federal, state, and local agencies. The public is expecting safer and more forgiving roadways and is demanding adequate compensation to victims. Tort claims can result in substantial monetary awards when deaths or disabling injuries are involved, even if the design and operation of the highway system are not the primary cause of accidents.

Utility poles located on highways' rights-of-way (R/Ws) have been identified as a major roadside hazard. Studies have estimated that between 2 and 5 percent of all accidents involve a utility pole (*1,2*). In 1990, highway accidents involving roadside hazards caused 12,780 deaths, of which 1,263 resulted from crashes with utility poles (*3*). Florida accounted for 10 percent of these fatalities. Pole-vehicle accidents in Florida caused 123 deaths, 6,195 injuries, and more than $24 million in vehicle damage in 1990 (*4*).

According to one study (*5*), a utility pole accident is 6 times more likely to result in a fatality and 3 times more likely to result in an injury than the average highway accident. Another study conducted at the University of Alabama (*6*) stated that because of the loss of the sovereign immunity status by most states, the number of tort claims against state highway agencies grew between 1978 and 1987 by an estimated 20 percent compounded annual rate to more than 27,000 in 1987. One study stated that utility poles had become frequent topics for tort liability actions (*7*). The University of Florida is conducting a study to assess the tort liability associated with utility pole accidents in Florida. The study is sponsored by the Florida Department of Transportation (FDOT) and FHWA. The study findings are presented next.

F. T. Najafi and F. E. Nassar, Department of Civil Engineering, University of Florida, 345 Weil Hall, Gainesville, Fla. 32611. P. Kaczorowski, Florida Department of Transportation, Office of Surveying and Mapping, 605 Suwannee Street, Tallahassee, Fla. 32301-8046.

FDOT'S TORT CLAIMS

Florida and other states have benefited from the increased federal and local funding for highway safety programs. It is estimated that in the past 15 years, the fatality rate on U.S. highways was reduced from 5.2 to 2.7 per million vehicle miles of travel (*8*). The reduction in pole-vehicle accidents in Florida is given in Table 1. However, unlike some DOTs that face a rapidly increasing liability cost, FDOT's general liability has been declining since 1988 as indicated in Table 2. FDOT's total losses from all claims and the premiums paid to the Department of Insurance (DOI) for the past 10 years are presented in Figure 1.

A cap of $100,000 per person and $200,000 per accident was imposed by a Florida statute on tort awards against state agencies. Tort suits against FDOT are handled by its Claim Office. Although the office handles several thousand claims every year, its files are not yet computerized. With the assistance of a legal secretary, claims related to utility pole accidents were sorted out. In addition to the files transferred to the Risk Management Office, only four additional claims were filed by accident victims that were related to utility poles. In the most recent case, the claimant sued FDOT on the basis of negligence in the installation, maintenance, and inspection of a utility pole. FDOT refused to pay damages because it has no maintenance responsibility. Another claim accused FDOT of not replacing a broken light bulb on a state road where the accident took place. FDOT rejected any responsibility because it was not notified of the broken lamp. The other two claims did not involve injuries. Plaintiffs accused FDOT of negligence and requested compensation for the repair costs of their vehicles. FDOT refused to accept responsibility. These four cases remain in litigation. However, no significant payments, if any, are expected.

The Risk Management Office handles tort claims against all state agencies. Because the office deals mainly with claims that result in payments, it uses a computerized data base for storing basic information about each claim. There were 19 cases involving a utility pole. All claims except three were indemnifications to utilities for damages caused by FDOT crews to utility poles. The largest indemnification was less than $2,000, and the average payment was $670. Only three claims were filed by individuals as a result of a pole accident, and none resulted in a monetary settlement.

TABLE 1 Utility Pole Accidents in Florida

YEAR	# ACCIDENTS	FATALITIES	INJURIES	VEHICLE DAMAGES
1989	8,355	181	7,254	$28,603,000
1990	6,873	123	6,195	$24,294,000
1991	6,211	128	5,733	$31,944,000

TABLE 2 FDOT's General Liability

YEAR	LOSSES	NUMBER OF CLAIMS	AVERAGE CLAIM
1991	$1,564,000	564	$2,772
1990	$2,629,000	942	$2,791
1989	$3,309,000	853	$3,879
1988	$4,668,000	1,962	$2,379
1987	$4,139,000	858	$4,824
TOTAL	$16,309,000	5,179	$3,149
AVERAGE	$3,262,000	1,036	$3,149

The Florida Public Service Commission (FPSC) has recently conducted a study on the cost-effectiveness of undergrounding electric utility wiring. FPSC located 115 sites of fatal pole accidents. Every site was visited, a descriptive form was completed, and several photographs were taken. These photographs, forms, and police reports were obtained and reviewed to estimate, somewhat subjectively, how many of these accidents could have been a good tort claim candidate. According to police reports, the principal cause of injury for about half of the accidents was excess speed. The other causes were loss of control at curves, impairment related to alcohol, collision with another vehicle, distraction, weather, or falling asleep. In addition, the photographs indicated that most poles had adequate lateral clearance. The few poles near the edge of the road were located in areas with limited R/W and would have been classified as special cases. Thus it is our opinion that very few accidents would have been a "good" tort case in Florida courts, although in other states, the courts have disagreed with the opinions of police officers and have found that DOTs or other parties contributed to accidents.

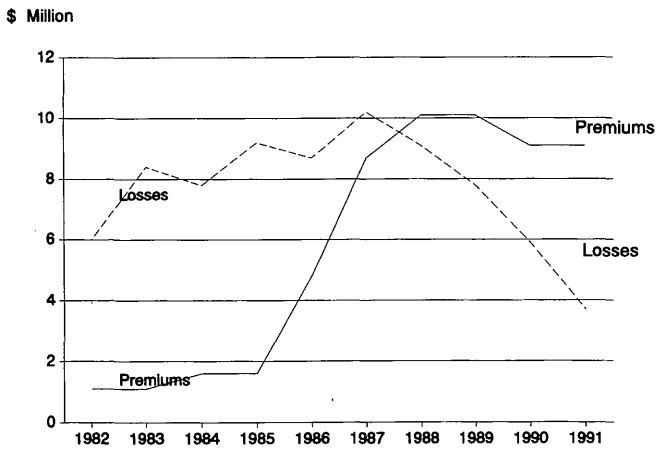


FIGURE 1 FDOT's premiums and losses, self-insurance (source: FDOT).

In summary, the review of FDOT and DOI tort files revealed that the tort liability related to pole accidents is not a problem for state agencies in Florida. This is the result of two factors: (*a*) the cap on awards and (*b*) favorable rulings by Florida courts as will be discussed later. There is, however, no guarantee that this favorable situation will continue into the future.

FLORIDA CITIES' TORT LIABILITY

A city can be named in a lawsuit if an accident happens on a road under its jurisdiction and the plaintiff has reasons to believe that the city did not design or maintain a safe road. Although the average speed is higher on state highways, resulting in more severe accidents, utility poles in urban areas are usually located closer to the traveling lanes.

The 20 most populous cities in Florida were identified as well as 5 smaller cities having a high rate of pole accidents. We requested information on completed tort cases related to pole accidents and conducted interviews with risk managers and safety engineers. We obtained the requested information from all surveyed cities except one major city. The survey's results are as follows:

1. More than half of the responding cities (mostly smaller cities) indicated that their records did not include any successful claim related to utility pole accidents. One large city paid more than $100,000 several years ago to settle a claim in which two persons were injured in a pole accident. This city successfully defended itself against other utility-related claims. The remaining cities had to deal with few claims. The

settlement amounts varied from a few thousand dollars to a maximum of $30,000 per case. None of these cities, however, considered the number of such claims excessive.

2. City officials did not regard the tort liability related to utility pole accidents as a major problem. On a scale of 1 to 10 (10 being a very serious problem), most officials gave a rank of 2 or 3. When asked to list priority highway safety improvements that would reduce tort claims against their city, relocating hazardous poles was generally not among cited improvements. Fixing pavement edges and broken sidewalks, improving intersection clearance, clearing obstructed Stop signs, trimming trees, relocating concrete poles used to support traffic signals, improving storm drainage curbing, and providing additional guardrails are a few of the priority safety improvements mentioned most often. Officials from cities that had to pay significant monetary awards to the parents of children killed or injured by falling from a bridge or injured by a culvert edge were very sensitive to similar cases and ranked these problems high on their priority list.

3. Most city officials consider the liability associated with utility poles to be a primary concern for utilities. They seemed confident that pole permits shifted liability to utilities responsible for maintaining the poles. City officials have to deal with more immediate problems where no one else shares responsibility. Many officials were surprised to learn about utility court cases in other states where DOTs and cities had to pay significant compensations.

UTILITIES' TORT LIABILITY

A one-page questionnaire requesting information on tort liability related to pole accidents was mailed to 63 utilities located in Florida (17 phone companies, 28 power companies, and 18 cable companies). More than half of the utilities responded to our survey. These companies can be divided into three groups: small utilities, public utilities, and large independently owned utilities. Most smaller utilities indicated that they were not subject to pay any court awards or other settlements to individuals injured by pole accidents. One company indicated that it paid less than $6,000 over a period of 3 years for eight out-of-court settlements. Many respondents indicated that they do not keep this kind of data. Most cable and small phone companies stated that most of their wiring is underground. Smaller utilities usually lease space on poles owned by large utilities.

Public utilities such as the Jacksonville Electric Authority and Gainesville Regional Utility benefit from the cap on tort awards against public agencies. Their records of completed tort cases are open to the public. Risk managers of public utilities indicated that the liability related to pole-vehicle accidents was not a major concern to them. Liability exposure due to environmental and health-related issues, electric contact accidents, and damages caused by storms and power failures were their main concerns. Their limited exposure and the favorable court ruling related to pole accidents in Florida helped keep this type of tort liability under control.

Southern Bell is the largest phone company in Florida. It owns 440,000 utility poles in Florida. The company informed us that for the past 5 years it was subject to 291 tort claims. Only nine of these claims were the result of pole accidents. The total monetary awards paid by Southern Bell was less than $60,000. Only one case is still being litigated. Regarding the liability related to joint use agreements of poles with other utilities, Southern Bell wrote: "Costs associated with liability and damages . . . have traditionally been shared with each party being liable for 1/2 of all injuries to any person or property. In recent years however, Southern Bell has taken the position that fifty percent liability is unfair considering the higher electric voltage in equipments used by electric companies. Consequently, recent agreements are void of language governing liability and damages and cases arising under these agreements would be decided by Florida Law."

The remaining phone companies, such as U.S. Sprint, MCI, Gulf Telephone Company, United Telephone Company of Florida, and Alltel Florida, and many television cable companies indicated that most of their cables and wires are underground. AT&T wrote that with the divestiture in 1984, 95 percent of utility poles went with the operating companies and AT&T exposure is minimal at best.

There are five independently owned electric companies (IOECs) in Florida. The IOECs are not required to report tort information to FDOT or to other public organization such as FPSC or other regulatory agencies. These companies regard tort information as very confidential and at first did not answer our survey. One agency wrote: "Plaintiffs in pole cases oftentimes seek to discover the same type of information. We regard such information as proprietary, privileged and confidential. We think that our successful defense of such requests would be jeopardized by responding to your survey." The IOECs must prepare two specific budgets every year: one for FPSC and the other for the Federal Energy Regulatory Commission (FERC). The IOECs report to FPSC on their payments for injuries and fatalities due to electric and nonelectric contacts. This reporting, however, is too general and not broken down by type of liability. Similarly, their reports to FERC include Sections 924 and 925, which deal with property insurance and injuries and damages, respectively, as indicated in Table 3.

Florida Power & Light (FPL) is the largest IOEC. The company's senior attorney elected not to provide specific information on tort liability. He said, however, that the liability related to pole-vehicle accidents is not a major concern for FPL. Environmental and health-related lawsuits are its primary liability concern. He also said that the favorable court decision in *Spiegel v. Southern Bell & Florida Power and Light* in 1977 has been upheld in Florida courts in recent cases. He wrote: "It is generally recognized that fault in connection with pole accidents lies with the driver who leaves the roadway out of control. Except in rare circumstances, the owners of structures, trees, or other objects located within the road rights-of-way are not deemed liable by Florida courts for damages or injuries to vehicles and their occupants who leave the travelled portion of the roadway and strike such objects. For that reason, we and other owners of properties and objects adjacent to roadways are rarely in the position of defending claims and lawsuits of this nature."

Tampa Electric Company is the third largest IOEC. Although the company did not answer our survey, we learned from an informed source that the company paid a significant amount to settle a 1985 claim. According to the same source, this seems to be the only case resulting in a large settlement.

TABLE 3 Independently Owned Utilities' Insurance Costs for 1990

UTILITY COMPANY	NUMBER OF EMPLOYEES	ELECTRIC OPERATION ($MILLION)	PROPERTY INSURANCE ($MILLION)	INJURIES & DAMAGES ($MILLION)
Florida Power and Light	15,500	$3,171	$16.30	$24.00
Florida Power Corporation	6,135	$1,050	$3.71	$4.77
Tampa Electric Company	3,218	$568	$1.92	$2.84
Gulf Power Company	1,615	$323	$1.78	$1.65
Florida Public Utility Company	73	$27	$0.06	$0.31

Gulf Power Company is the fourth-largest IOEC and owns about 400,000 poles. Gulf Power was sued jointly with FDOT by a driver who hit one of its poles (*Kay v. Gulf Power*). A circuit jury returned $7 million compensatory damages against both defendants and $4.2 million punitive damages against Gulf Power only. This verdict was later reversed and remanded by the First District Appellate Court in 1986. This case is discussed in a later section of the paper. After the case rehearing was denied, Gulf Power settled with the plaintiff out of court. In February 1988, Gulf Power settled another case. Although we were given the settlement amounts for both cases, we were requested not to publish them. We were also told that these were the only pole-related cases settled by Gulf Power in recent years.

In summary, tort liability related to pole-vehicle accidents is not a problem for phone and cable companies. However, the situation is unclear for IOEC. This favorable situation is due primarily to sustained favorable court rulings. Privately owned utilities can be subject to significantly more expensive claims if Florida courts become more sympathetic to accident victims. Although this happened in other states such as California, utilities' officials seemed confident that the legal situation will not change significantly in Florida.

INSURANCE COMPANIES' RESPONSES TO THE SURVEY

A one-page questionnaire was mailed to the 50 largest insurance companies. Twenty-five responded, but none completed the questionnaire. Although most respondents commended the university for the study and many requested a copy of the study's final report, some of their representative key statements are as follows: "We are currently unable to retrieve such specific information"; "the information requested is not captured on any of our automated systems"; "we do not have the ability to break down the information as requested"; "we have no way of extracting the information you seek." Some insurance companies simply indicated that they do not insure utility poles. State Farm Insurance Company, the nation's largest insurance company, stated that "we are unable to identify, from electronically encoded records, accidents involving collision with highway poles. In any case, State Farm auto liability insurance would not pay on behalf of the owner of a pole for liability arising out of its faulty design or placement."

Large utilities in Florida are self-insured to a certain limit (e.g., $1 million in the case of Gulf Power), and they buy layers of additional coverage from insurance companies. A blanket coverage is bought on the basis of historical and expected claims. Premiums are based on basic information about the company and its location and are adjusted every year to reflect a change in liability payments. Premiums, however, are not broken down by the type or number of utility poles. One national insurance company, AEGIS Insurance Services, specializes exclusively in insuring utilities. AEGIS was reluctant to provide us with any specific tort information about its clients. It later indicated that it will reconsider our request, but we still have not received any information.

TORT CLAIMS IN OTHER STATES

A one-page questionnaire was mailed to transportation agencies and risk management offices in other states. Most of the states completed the questionnaire. About half of the respondents stated that the utility permit's liability clause shifted poles' entire liability to utilities, and therefore they paid no compensations to pole accident victims. Pennsylvania paid more than $0.5 million to settle 25 pole cases. The other respondents indicated that neither their insurance commissioner office nor their DOT compiles such information. About 20 percent of respondents mailed a copy of their state's utility permit. Most permits include a clause containing a sentence similar to the following: "Permittee hereby agrees to indemnify and hold harmless the State . . ."

A questionnaire was also mailed to 50 utilities located outside Florida. Only 10 utilities responded. Very little tort information was provided, except for one utility in California, which provided us with confidential information showing substantial settlements and court awards and a large number of unresolved cases. Clearly, the tort liability associated with utility poles differs among states.

Information was also requested from organizations representing private and public interests such as Public Risk Management Association, Insurance Institute for Highway Safety, Insurance Information Institute, and Edison Electric Institute. All of these organizations except Edison Electric Insti-

tute indicated that they do not collect the type of information we are seeking. Edison Electric Institute has decided not to participate in this study.

COURT RULINGS IN FLORIDA

The outcome of utility pole tort liability rulings depends on two main factors: (*a*) the state's immunity status and the maximum cap on monetary awards and (*b*) the utility permit's liability clause and utility pole design guidelines.

The state immunity status is enacted by the state legislative body and approved by the state supreme court as constitutional. Immunity status can be amended at any time by legislative action or after the state supreme court declares it to be no longer constitutional.

Each state has formulated its own utility permit document. Although most states use similar statements in their document's liability clause, a small change in wording can have a significant legal consequence. For instance, FDOT amended in 1989 its permit's liability clause. The amendment was challenged in court by utilities and finally struck down by the Florida Supreme Court.

Utility pole design criteria are related to the state's clear zone policy. Florida's *Utility Accommodation Guide,* published in 1990, provides design criteria for highways' clear zone and utility poles located on highways' R/Ws. In locations with limited R/Ws, these guidelines allow special cases and call for "reasonable judgments," which may be interpreted differently in courts. In general, negligence and nuisance are the leading basis for tort liability. A review of key court rulings in Florida is presented next.

Immunity

State agencies in Florida, including FDOT, benefit from a limited immunity. FDOT has immunity in planning or discretionary decisions but has lost its immunity in operational or proprietary decisions through the "waiver of immunity" statute. A state statute limiting monetary awards to $100,000 per individual was enacted. Planning-level decisions are related to such matters as location and placement of traffic and pedestrian control devices (*Perez v. FDOT,* 1982; *Lewis School v. Metropolitan Dade County,* 1979; *Gordon v. West Palm Beach,* 1975), decisions on a road extension or on guardrail placement (*Payne v. Palm Beach County,* 1981; *Hyde v. FDOT,* 1984), and decisions on setting speed limits (*Ferla v. Metropolitan Dade County,* 1979), among others.

Operational-level decisions are primarily related to the operation and maintenance of roadways and shoulders (*Wojtan v. Hernando County,* 1980; *Trumpe v. Coral Spring,* 1976). Negligence in design or construction of a facility is also an operational decision (*Gordon v. West Palm Beach,* 1975). Failure to maintain a stop sign is an operational decision (*Wallace v. National Mutual,* 1979). However, unless negligence is proven in failing to repair a damaged or malfunctioning traffic signal, the defendant is usually not liable (*Arenado v. FPL,* 1988).

A court decision in 1982 ruled that the location of individual poles and guy wires to support power lines was an operational function falling outside governmental tort immunity and was therefore subject to liability on the basis of negligence (*Austin v. City of Mt. Dora,* 1982). Two subsequent court rulings disagreed with this decision. In the case of *Miller v. Fort Lauderdale,* the court decided in 1987 that the location of a street light pole was a planning-level decision that is immune from liability. In 1988, the court also decided in the case of *Hosey v. Fort Lauderdale* that the placement of a street light pole on a divided island was a discretionary planning-level decision and the city was shielded from suit.

Negligence

To recover against FDOT for injuries, it must be shown that the department was negligent for creating or knowing about a dangerous condition and failing to correct such condition or warn the public about it (*Hodges v. WinterPark,* 1983). Governmental entities have a duty to exercise reasonable care to maintain highways in a safe condition (*Tamarac City v. Garchar,* 1981). Utilities have a duty to exercise care, both in location or construction and in use and maintenance of their lines, poles, and equipment. They must do all that human care, vigilance, and foresight can do to protect the safety of the public. However, they are not under continuing duty to protect against unforeseeable or unlikely events (*Padgett v. West Florida Electric Company,* 1982). In the 1977 *Spiegel v. Southern Bell & FPL* case, plaintiff's attorneys alleged that the companies negligently maintained a pole so near the highway that driver was fatally injured when his vehicle collided with the pole. The Circuit Court for Dade County rendered a final judgment in favor of defendants. Plaintiff's attorneys appealed. However, the District Court of Appeal held that the trial court was correct in determining that the utilities were not liable. In the case of *Gulf Power Company v. Kay,* the District Court of Appeal found that DOT's guides and manuals were not applicable to Gulf Power at the time of accident, and the company knowledge of two previous car accidents hitting the same pole was not legally sufficient to warrant punitive damages, because of the lack of similarity of circumstances between accidents.

Punitive Damages

The Florida Supreme Court stated that the character of negligence necessary to sustain an award of punitive damages must be of a gross and flagrant character, evincing reckless disregard of human life, which would raise presumption of a conscious indifference to consequences or a gross disregard of the safety and welfare of the public. On this basis, the jury award of $4.2 million in punitive damage against Gulf Power (*Kay v. Gulf Power*) by a lower court was reversed by the Appellate Court.

Design Criteria and Standards

A change to a code governing construction standards is not given retroactive effect to construction completed before its adoption, except under limited circumstances. That is why, in the case of *Gulf Power v. Kay,* DOT design standards were not admissible in court—they were not applicable at the time

of construction. However, these guides are applicable to reconstruction or major improvement projects. FDOT and utilities are liable for failing to comply with statutory standards and criteria for design, construction, and maintenance of public roads (*FDOT v. Neilson,* 1982; *Ferla v. Metropolitan Dade County,* 1979).

CONCLUSIONS AND RECOMMENDATIONS

Utility poles are a major roadside hazard for errant motorists who leave the roadway. They account for about 15 percent of accidents with fixed objects and cause many injuries and fatalities. Although tort claims against state highway agencies are increasing at a fast rate because of loss of immunity status and higher caps on awards, it was found that the situation in Florida did not conform to the national trend. No significant payments were made by the state to settle tort claims related to pole accidents. Furthermore, the state's indemnification to utilities for shared liability is almost nonexistent. City officials indicated that the tort liability resulting from pole accidents was generally not considered a serious problem. As for utilities, the situation is unclear. The tort information provided by independently owned utilities indicated that tort liability associated with utility poles is not a primary concern for them. Although Florida courts rendered generally favorable rulings in favor of utilities, utilities paid in a few cases substantial out-of-court settlements.

This generally favorable situation for public agencies and utilities in Florida is unlikely to change drastically in the near future. However, a long-term assessment of the situation may provide a different picture for the following reasons:

1. Poles will continue to account for a disproportionately high rate of accidents and injuries.
2. Many utility poles are located near the traveling lanes, especially in urban areas. Courts may be more sympathetic to drivers or passengers severely injured by colliding with such a pole even if these poles satisfy design standards.
3. The cap on awards against public agencies is likely to increase as is the case in other states, which can result in more claims.
4. Utilities' successful court challenge of Florida's amendment of its utility permit's liability clause indicates that future permit clauses may increasingly reflect utilities' interests.
5. The emphasis on highway safety will increase public expectation regarding the elimination of highway hazards.

The confidence in continuing favorable court rulings expressed by many city and utility officials cannot be justified in view of the national legal trend. It only takes one sustained court ruling sympathetic to drivers injured or killed by a pole accident to create a legal precedent and affect all subsequent cases. Therefore, it is prudent for utilities and highway agencies to institute new policies to deal with the tort liability risk related to pole accidents. This policy may include the following:

1. Relying on computerized data bases to store, retrieve, and evaluate tort claim data;
2. Developing a uniform method based on a well-structured coding system to collect and store tort information;
3. Requiring utilities to provide more detailed information on tort liability in their annual reports to regulatory agencies;
4. Standardizing the liability clause related to poles' joint use contracts among utilities;
5. Formulating criteria for an equitable sharing of relocation cost between FDOT, cities, and utilities when pole relocation is necessary as a preventive measure against tort claims;
6. Establishing an independent organization to provide national and state summaries of court rulings and jury awards in related cases; and
7. Developing a method to prioritize the relocation of existing hazardous poles. This method should preferably be based on an expert system model, because expert systems rely on rules to deal with nonanalytical problems and thus can incorporate experts' knowledge. An effective expert system program is capable of learning from every new case and automatically updating its rules and adjusting relocation priorities on the basis of new court rulings. However, this requires uniform and "automated" data reporting methods adopted by all parties.

Pole relocation priorities should be based on tort information and court ruling interpretations instead of on accident statistics. The reason is that accident frequency and liability cost do not necessarily depend on the same factors.

REFERENCES

1. *Selection of Cost-Effective Countermeasures for Utility Pole Accidents.* Users Manual. Report FHWA-IP-86-9. FHWA, U.S. Department of Transportation, Dec. 1986.
2. I. S. Jones and A. S. Baum. *An Analysis of the Urban Utility Pole Problem.* Calspan Field Services, Inc., Federal Highway Administration, Dec. 1980.
3. *Roadside Hazards, Fatality Facts 1991.* Insurance Institute for Highway Safety, Arlington, Va., July 1991.
4. *Report on Cost-Effectiveness of Underground Electric Distribution Facilities.* Florida Public Service Commission, Tallahassee, Fla., Dec. 1991.
5. G. B. Pilkington II. Utility Poles: A Highway Safety Problem. *Public Roads,* Vol. 52, No. 3, Dec. 1988.
6. D. S. Turner, J. K. Davis, and B. T. Wood. *Transportation Research Circular 361: Status Report: Tort Liability Among State Highway Agencies.* TRB, National Research Council, Washington, D.C., July 1990.
7. D. S. Turner. A Primer on the Clear Zone. In *Transportation Research Record 1122,* TRB, National Research Council, Washington, D.C., 1987, pp. 86–95.
8. J. A. Cirillo and F. M. Council. Highway Safety: Twenty Years Later. In *Transportation Research Record 1068,* TRB, National Research Council, Washington, D.C., 1986, pp. 90–95.

Publication of this paper sponsored by Committee on Tort Liability and Risk Management.

Hydroplaning and Roadway Tort Liability

JOHN M. MOUNCE AND RICHARD T. BARTOSKEWITZ

Tort claims against highway agencies for alleged incidents of hydroplaning due to roadway defects have been growing in number. Many claims of hydroplaning cannot be substantiated by the weather, roadway, or vehicle conditions present at the time of the accident. And often when hydroplaning occurs, the evidence indicates that an inappropriate response to adverse driving conditions, or simply driver negligence, may be the direct cause rather than a roadway defect. Research of the phenomenon of hydroplaning was reviewed to address issues that arise when hydroplaning is alleged in roadway tort litigation. Hydroplaning is the separation of a rolling or sliding tire from the roadway surface by a layer of fluid. Of the three types of hydroplaning commonly recognized, highway engineers are primarily concerned with viscous and dynamic hydroplaning. Of these two, dynamic hydroplaning presents the greater risk. In the extreme situation of full dynamic hydroplaning, complete separation of the tire from the pavement by a fluid layer negates the driver's ability to control vehicle speed and direction. Hydroplaning may be avoided by consideration of several factors. Proper highway design may reduce hydroplaning risks by providing adequate pavement texture and cross slope. However, ultimate responsibility for avoiding hydroplaning lies with the driver. Drivers can reduce incidents of hydroplaning by maintaining tires in good condition at rated inflation pressures and by slowing down during rainstorms or on wet roadways.

Rainfall and water present on the pavement surface influence the safety of motor vehicle operation. The latest national accident statistics, compiled through 1990, indicate that approximately 10 percent of all fatal crashes occur on wet pavements during rainfall (*1*). In Texas, approximately 28 percent of all accidents are categorized as occurring during rainfall or on wet pavements (*2*).

Motorists must be relied upon to recognize the degradation of their ability to operate safely brought on by diminished visibility through rainfall and reduced friction capabilities on wet pavement. Many accidents in wet weather are due to loss of vehicle control, which results from either failure to recognize or to properly respond to adverse weather and pavement conditions.

In recent years, an increasing number of tort lawsuits have been filed against street and highway operating agencies with allegations of roadway defects responsible for "hydroplaning." In the adjudication of these lawsuits, many statements have been made as to when, where, and how hydroplaning occurs. Most wet weather accidents are not caused by hydroplaning. In reality, hydroplaning is a rare event, and its occurrence is dependent on many factors. This paper is a compilation of research directed to the phenomenon of hydroplaning as related to roadway tort litigation.

Texas Transportation Institute, The Texas A&M University System, College Station, Tex. 77843-3135.

PHYSICS OF HYDROPLANING

A basic understanding of the function of pavement texture in the tire-pavement interface is critical to a discussion of the mechanics of hydroplaning. Roadway surfaces are characterized by pavement microtexture and macrotexture. Microtexture describes the degree of polishing of the pavement surface or aggregate, varying from harsh to polished (*3*, Chapter 2), and is necessary to the development of frictional forces between the tire and pavement on wet surfaces. The magnitude of these frictional forces becomes greater with increased microtexture, and it is maximized at lower vehicle speeds (*4*). When a thin layer of water is present, asperities on the pavement surface break through the waterfilm to enable direct contact between the tire and pavement (*5*). These asperities are thousands of small, pointed projections that make up microtexture. High local bearing pressures are generated by contact between the tire tread and the pavement asperities, thereby allowing the tire to establish essentially "dry" contact with the roadway (*6*).

Macrotexture describes the size and extent of large-scale protrusions from the surface of the pavement, varying from smooth to rough. Macrotexture is a function of aggregate gradation, the pavement construction method, and special surface treatments such as grooving or chipping (*3*, Chapter 2). Whereas microtexture governs wet friction at low vehicle speeds, macrotexture is the critical factor for higher vehicle speeds. Friction levels are significantly lower for pavements with poor macrotexture than for pavements with good macrotexture when vehicle speeds are high and flooded conditions prevail. This is explained by the fact that macrotexture provides channels for drainage, thereby reducing hydrodynamic pressures existing between the tire and pavement when water is present (*4*). For a thin waterfilm and high vehicle speeds, macrotexture is vital to establishing and maintaining contact between the tire and pavement. For a flooded pavement, it operates as escape channels for bulk water drainage from beneath the tire footprint (*6*).

The physical phenomenon of hydroplaning is the separation of a rolling or sliding tire from the roadway surface by a layer of fluid. On a wet or flooded pavement, hydrodynamic pressures increase as vehicle speed increases and eventually reach a critical point at which the tire is lifted away from the surface (*7*). Three types of hydroplaning have been identified: (*a*) viscous hydroplaning, (*b*) dynamic hydroplaning, and (*c*) tread

rubber reversion hydroplaning. Viscous and dynamic hydroplaning are of concern when examining highway operations on wet pavements.

Viscous hydroplaning is a problem associated with low-speed operation on pavements with little or no microtexture. It results from an extremely thin film of water existing cohesively between the tire and the pavement surface because of insufficient microtexture to penetrate and diffuse the fluid layer. For this reason, viscous hydroplaning is commonly referred to as thin film hydroplaning to distinguish it from dynamic hydroplaning, which requires a comparatively thick fluid layer.

Opinions on the importance of vehicle speed to viscous hydroplaning vary. Yeager states that viscous hydroplaning is observed at vehicle speeds greater than 32 km/hr (20 mph) (*8*). However, Browne contends that viscous hydroplaning can occur at any vehicle speed and with any waterfilm thickness (*9*). The important point is that it may occur when vehicle speeds are very low, such as with speeds typical of city driving. The most critical factors of influence during viscous hydroplaning are the viscosity of the fluid, tire condition, and the quality of the pavement surface. It will not occur unless the tire tread depth is very shallow and the pavement has a "polished" quality. Viscous hydroplaning may be described as a rare event characterized by a bald tire operating on a mirror-smooth surface.

Dynamic hydroplaning results from uplift forces created by a water wedge driven between a moving tire and the pavement surface, as shown in Figure 1. The risk of dynamic hydroplaning is high when fluid inertial effects dominate, as with thick waterfilms found on a flooded pavement. Dynamic hydroplaning can only occur when the water accumulation encountered by the tire exceeds the combined drainage capacity of the tire tread and the pavement macrotexture for a given speed (*9*). For extreme conditions, it has been observed for water depths as little as 0.76 mm (0.03 in.) with bald tires on smooth, polished pavement surfaces (*8*).

A hydroplaning tire may experience either partial or full dynamic hydroplaning. With partial dynamic hydroplaning, only part of the tire actually rides on the surface of the water. Contact between at least a portion of the tire footprint and the pavement surface is maintained. Full dynamic hydroplaning, on the other hand, is characterized by complete separation of the tire from the pavement by the fluid layer. The occurrence of full dynamic hydroplaning represents a far greater hazard than partial dynamic hydroplaning; the driver is unable to control vehicle steering and braking because of the loss of contact.

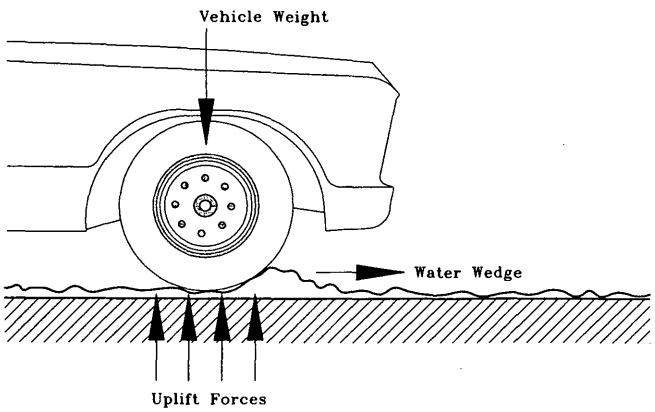


FIGURE 1 Dynamic hydroplaning.

Speed and waterfilm thickness are the governing conditions for partial and full dynamic hydroplaning. It is difficult to identify with precision the speed at which these phenomena occur, because other variables that describe the roadway surface, the tire condition, and the driving environment must be considered. Whereas ordinary highway operating speeds and water depths may give rise to partial dynamic hydroplaning, considerably higher vehicle speeds and a very thick waterfilm, such as that produced by high-intensity rainfall, are necessary for full dynamic hydroplaning to occur (*10*). For most situations, the vehicle speed at which full dynamic hydroplaning is observed would be considered unsafe or not prudent for the amount of water on the roadway, assuming that the tire tread is sufficient and that the tires are properly inflated.

FACTORS INFLUENCING ROADWAY HYDROPLANING

Dynamic hydroplaning is a function of the complex interaction between many variables. For this reason, the probability of full dynamic hydroplaning is rather low (*10*). Factors critical to hydroplaning are shown in Figure 2. As can be seen, the four primary effective variables are rainfall, the roadway, tire characteristics, and the driver.

In general, hydroplaning is a low-probability event because rainfall intensities necessary to flood a pavement surface are rare and of short duration (*11*). Furthermore, rainfall intensities of sufficient magnitude [5.1 to 10.2 cm/hr (2 to 4 in./hr)] to create sheet flooding of pavement surfaces reduce visibility even with wipers so that prudent drivers will reduce operating speeds for safety (*10*).

Drainage path length refers to the distance any discrete water molecule would have to negotiate to drain from a given point on the pavement surface. It is a function of the number of lanes of travel and the lane width. A typical two-lane,

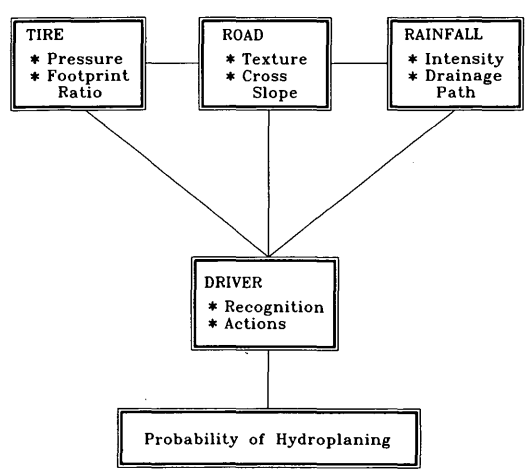


FIGURE 2 Synthesis of interactive factors influencing hydroplaning.

crowned cross section has a nominal drainage path length of 3.66 ft. This factor is especially significant for water accumulations that result from extended drainage path lengths associated with multilane roadways. An investigation of potential means of decreasing the occurrence of hydroplaning concluded that minimizing the drainage path length through careful highway design and construction is an effective strategy (*11*). When multiple travel lanes are present, the negative impact of longer drainage path lengths can be mitigated through appropriate application of pavement cross slope and pavement texture.

Roadway factors of pavement texture and transverse cross slope are critical to controlling water accumulation and drainage. A transverse cross slope of 2.5 percent is desirable to facilitate adequate surface drainage for common rainfall intensities without impeding vehicle steering or lane-changing maneuvers (*10*).

The role of pavement texture in collecting and draining surface water from the vehicle path has already been addressed. Balmer and Gallaway (*11*) reported the results of an extensive investigation of applications of pavement texture to reduce the risk of hydroplaning and to improve wet traction. The use of a gritty, coarse surface texture or finish in the construction and maintenance of pavements was recommended.

Providing texture depth is also critical because deeper textures act as larger escape channels for water forced from beneath the tire footprint region. Balmer and Gallaway discovered that increasing the texture depth from 0.76 mm (0.03 in.) to 3.81 mm (0.15 in.) raised the speed at which dynamic hydroplaning was predicted to occur by 16.1 km/hr (10 mph) for tire inflation pressure of 206.85 kPa (30 psi), tire tread depth of 6.75 mm (8.5/32 in.), and water depth of 7.6 mm (0.3 in.). It was also concluded that transverse texture, aligned parallel to the cross slope direction, can be expected to provide improved overall surface drainage, improved water expulsion between the tire and the pavement, and a decrease in the forward motion of water responsible for creating a water wedge between the tire and pavement.

Pavement texture depth of 1.52 mm (0.06 in.) or greater is the recommended minimum for roadways with high operating speeds. This will provide adequate drainage and decrease hydroplaning for normally expected rainfall rates (*10*). For roadways with low-speed operation, even less texture depth may be tolerable. However, even under the best design and construction conditions, storms of unusually high intensity, though rare, are likely to create flooding of the pavement surface above the texture asperations.

The tire is one of the most critical factors influencing hydroplaning. Even on a well-designed, properly maintained roadway, a worn, under-, or overinflated tire experiences considerably higher risk of hydroplaning than does a tire in "good" shape, for normally expected rainfall and prudent speed. Yeager (*8*) and Browne (*9*) have addressed factors of tire construction and condition that influence hydroplaning.

Tread pattern is one of these factors. Lateral and longitudinal grooves, sipes, and ribs make up the tire tread pattern. Grooves are the deep channels that run around the circumference of the tire (longitudinal grooves) and across the tire surface (lateral grooves). They serve two principal functions. By channeling bulk water through and out of the tire footprint region, grooves help to prevent the formation of the water wedge that penetrates into the footprint region and causes dynamic hydroplaning. They also function as reservoirs for thin waterfilms squeezed from between the tire and the pavement surface, which reduces the risk of viscous hydroplaning (*9*).

Four parameters describe the effectiveness of the tread grooves with respect to wet traction and hydroplaning: tread depth, groove capacity, groove shape, and groove spacing. Tread depth is primarily a measure of how much tread remains on a tire after experiencing wear as a result of extended use. When the tire is worn to an extent such that the depth of tread reaches a minimum safe value, tire replacement is recommended.

The amount of surface water to be effectively handled is referred to as the tire's groove capacity. It is related to tread depth and influenced by tire construction, load, and inflation pressure. Once the amount of pavement surface water encountered by the tire tread exceeds the groove capacity, the excess water must have sufficient time to be displaced without building up in front of the tire and creating uplift pressure on the tire. Higher vehicle speeds reduce the time of displacement and increase the risk of hydroplaning.

Another determining factor of groove capacity is groove closure. The effect of groove closure is a considerable reduction in the tread's groove capacity. This phenomenon depends on the structural properties of the tire tread, the rotational speed of the tire, and the inertial forces of the fluid layer that the tire encounters. It is a direct consequence of lateral forces acting in the tire ribs toward the longitudinal centerline of the tire footprint. Groove closure is resisted by frictional forces between the tire and the pavement. However, in the absence of these frictional forces, such as on a wet pavement, no force exists to counteract groove closure. Groove closure has been found to be less of a problem for radial tires than for bias tires (*8*).

Groove shape and spacing influence a tire's wet traction capabilities and performance. Groove shape is especially important for a sliding tire, as opposed to a free rolling tire (*8*). Wide grooves provide optimum flow characteristics and mitigate the effects of groove closure. Slight amounts of zigzag with diagonal grooves are also desirable. For a free rolling tire, groove capacity is the controlling factor, although diagonal grooves and blading help to reduce the risk of viscous hydroplaning on a smooth surface. Grooves should be closely spaced to achieve peak traction performance.

Other tire factors relating to hydroplaning and wet traction may be generally categorized as elements of the tire carcass. These include tire dimensions and flexibility. The region of contact between the tire and pavement, the tire footprint, is measured by length and tire width. As tire width increases, the width of the footprint increases. On a wet or flooded pavement, this is important because the tire will encounter and interact with a greater amount of fluid than it would have otherwise. Accordingly, the task of collecting and channeling water away from the tire footprint becomes more difficult and requires a greater length of time, and the magnitude of hydrodynamic forces acting on the tire is greater. But whereas increasing the width of the contact region is potentially detrimental, increasing its length results in greater amounts of dry contact within this region. It follows that wet traction performance and safety are enhanced.

The effects of tire footprint dimensions on dynamic hydroplaning speed have recently been investigated (*12–14*). The tire footprint aspect ratio is calculated as the tread contact area width divided by the length of the footprint (Figure 3). It is of particular interest in analyzing the hydroplaning tendency of tractor-trailer trucks. Aspect ratios for trucks are influenced by the magnitude of the load. The footprint aspect ratio for an empty truck is considerably higher than for a loaded truck, when holding inflation pressure constant, due to shorter tire footprints for empty trucks. As explained previously, this results in less dry contact area between the tire and the pavement. Furthermore, accident statistics show that jackknifing of empty tractor-trailer trucks on wet pavements is a significant event that may be attributed to dynamic hydroplaning. It was determined that the footprint aspect ratio is a variable that must be considered when estimating dynamic hydroplaning speeds for pneumatic tires.

Tire construction and inflation pressure govern tire flexibility. Bias ply, belted bias ply, and radial ply are the three common methods of tire construction. With respect to decreasing the potential of the tire to hydroplane, belted bias ply and radial tires are preferred. The treads of these tires have improved stability, provided by belts under the tread region. This serves to reduce tire tread wear and groove closure and makes possible the inclusion of exaggerated tread patterns that reduce hydroplaning risks (*9*).

The function of tire inflation pressure in raising or lowering a tire's hydroplaning tendency is difficult to analyze and evaluate. It has been shown that for dynamic hydroplaning to take place, the tire surface must deform inward, toward the center of the tire. When this deformation is present, water can penetrate deeper into the tire footprint to create the water wedge that can eventually lead to full dynamic hydroplaning. Higher inflation pressure improves the tire's rigidity and its ability to resist the hydrodynamic forces causing tire surface deformation, thereby raising the speed required for hydroplaning to occur. It also counteracts the lateral forces in the tire ribs that encourage groove closure. The drawback, however, is shortening of the tire footprint and the ensuing reduction of the dry contact area between the tire and pavement. This essentially lowers the hydroplaning speed (*9*).

Roadway, vehicle, and environmental factors that interact to create hydroplaning have been mentioned. The driver's recognition of and response to these various factors are critical. Drivers avert hydroplaning by direct action, for instance by maintaining safe speeds on wet roadways. They can also indirectly reduce the potential for hydroplaning through a careful program of tire maintenance.

PREDICTING AND IDENTIFYING HYDROPLANING SPEEDS

Substantial effort has been devoted to the development of formulas and criteria to identify the precise speed at which hydroplaning occurs. The most common approach has been to calculate the critical speed required for dynamic hydroplaning. Some of these equations are simple relationships defining the hydroplaning speed as a function of one or two variables. Others are considerably more complex. As might be expected, the task of predicting when hydroplaning will occur, or of identifying a particular wet-weather accident as a hydroplaning incident, is rather difficult and involves a substantial degree of uncertainty. The purpose of this section is to briefly describe some of the analytical and empirical techniques for evaluating hydroplaning potential.

In the case of viscous hydroplaning, Equation 1 describes the minimum hydroplaning speed for a pavement surface with slight microtexture:

$$V_H \geq \frac{L}{\Delta T_{sf}} \tag{1}$$

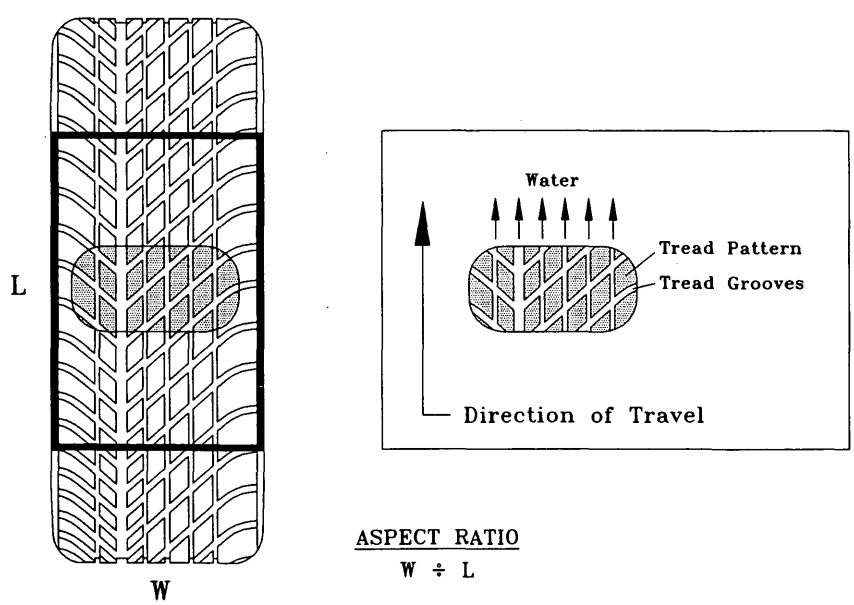


FIGURE 3 Tire footprint—pavement view.

where

V_H = minimum viscous hydroplaning speed,
L = length of the tire footprint region, and
ΔT_{sf} = time required for sufficient reduction of the fluid film for contact between the tread rubber and the pavement asperities to occur (*9*).

This formula is not applicable to dynamic hydroplaning.

Yang has proposed an analytical equation to define hydroplaning as part of an effort to develop design criteria for runway pavement grooving (*15*). The underlying principle for this equation is that hydroplaning will occur when the water escape velocity due to an external force, the tire pressure, is less than the speed at which the surface water travels sideways. The critical moment at which hydroplaning occurs is defined by Equation 2:

$$cp^{1/2} = 0.1292\left(\frac{\pi a v}{b}\right) \tag{2}$$

where

c = constant,
p = tire inflation pressure (kPa),
a = width of the tire footprint (cm),
b = length of tire footprint (cm), and
v = vehicle velocity (cm/sec).

For U.S. customary units, Equation 2 is rewritten as

$$cp^{1/2} = \frac{\pi a/4}{2b/v} \tag{3}$$

where

c = constant,
p = tire inflation pressure (lbf/in.^{2}),
a = width of the tire footprint region (in.),
b = length of the tire footprint region (in.), and
v = vehicle velocity (in./sec).

The development of this equation assumes an elliptical tire footprint shape.

One of the most frequently cited hydroplaning equations was developed by Horne to predict the minimum dynamic hydroplaning speed for pneumatic tires (*16*). In its simplified form, this equation is

$$V_H = 6.35\sqrt{p} \tag{4}$$

which yields the minimum tire hydroplaning speed V_H (km/hr) as a function of the tire inflation pressure p (kPa). In U.S. customary units, Equation 4 is given by

$$V_p = 10.35\sqrt{p} \tag{5}$$

where the minimum tire hydroplaning speed V_P is in mph and the tire inflation pressure p is in lbf/in.^{2} The formula is derived from empirical data and based on inertial properties of the fluid layer. It is applicable to flooded pavements, when the water depth exceeds the tire tread depth.

Recent research has indicated that the minimum dynamic hydroplaning speed of automobile, truck, and bus tires varies not only with the inflation pressure but also with the tire footprint aspect ratio (*12–14*). Consequently, Horne proposed a modification of his earlier formula to account for the influence of the footprint aspect ratio under load. Simplified, this new equation may be written as Equation 6:

$$V_H = 4.87\sqrt{p(w/l)^{-1}} \tag{6}$$

where w/l is the tire footprint aspect ratio, the tire inflation pressure p is in kPa, and the minimum tire hydroplaning speed V_H is in km/hr. For U.S. customary units, Equation 6 may be written as

$$V_p = 7.95\sqrt{p(w/l)^{-1}} \tag{7}$$

which yields the speed V_P in mph as a function of the tire inflation pressure p in lbf/in.^{2} It is seen that the magnitude of the minimum dynamic hydroplaning speed increases as the tire inflation pressure increases and the tire footprint aspect ratio decreases (*12*). Research at the Texas Transportation Institute (TTI) investigated the validity of Horne's predictions of dynamic hydroplaning of lightly loaded truck tires at typical highway speeds (*13*). TTI engineers formulated the relationship

$$V = 24.99(p)^{0.21}\left(\frac{1.4}{w/l}\right)^{0.5} \tag{8}$$

normalized for the test aspect ratio of 1.4. In U.S. customary units, Equation 8 is written as

$$V = 23.3(p)^{0.21}\left(\frac{1.4}{w/l}\right)^{0.5} \tag{9}$$

Although Equations 8 and 9 differ from Equations 6 and 7, they yield curves that agree closely over the range of test conditions.

A study by Gallaway et al. (*17*) developed an empirical formula for dynamic hydroplaning speed when the waterfilm thickness exceeds 0.10 in. Multiple linear regression yielded the following expression:

$$V = 0.902SD^{0.04}P^{0.3}\left(\frac{TD}{0.794} + 1\right)^{0.06} A \tag{10}$$

where A is the greater of

$$A = \left(\frac{11.008}{WD^{0.06}} + 3.507\right) \tag{11}$$

or

$$A = \left(\frac{26.871}{WD^{0.06}} - 6.861\right)TXD^{0.14} \tag{12}$$

and V is the vehicle speed (km/hr), SD is the spindown percentage, P is the tire inflation pressure (kPa), TD is the tread depth (mm), WD is the water depth above the pavement asperities (cm), and TXD is the pavement texture depth (cm).

To indicate the point at which hydroplaning occurs, the spindown parameter was used. Spindown describes the change in a free rolling tire's rotational velocity upon loss of contact with the pavement surface, as in full dynamic hydroplaning. When U.S. customary units are used, Equation 13 is applied:

$$V = SD^{0.04} P^{0.3} (TD + 1)^{0.06} A \quad (13)$$

where A is the greater of

$$A = \left(\frac{10.409}{WD^{0.06}} + 3.507\right) \quad (14)$$

or

$$A = \left[\frac{28.952}{WD^{0.06}} - 7.817\right] TXD^{0.14} \quad (15)$$

and V is expressed in mph, P is in lbf/in.^{2}, TD is given as 32nds of an inch, and WD and TXD are expressed in inches.

Two studies conducted at The Pennsylvania State University have investigated hydroplaning speeds. Agrawal et al. (*18*) ranked highway pavement performance by evaluating the hydroplaning potential of various pavement treatments. The dynamic hydroplaning speed was determined indirectly by measuring the brake force coefficient, the friction value that describes the tire-pavement interface. It was assumed that full dynamic hydroplaning occurs when the brake force coefficient is zero.

Huebner et al. (*19*) developed a hydroplaning model that draws on the work of both Gallaway and Agrawal. For waterfilm thicknesses greater than 0.25 mm (0.10 in.), Gallaway's equation for the critical dynamic hydroplaning speed was adopted. A regression of 18 data points collected by the Agrawal study for waterfilm thicknesses less than 0.25 cm (0.10 in.) was performed. The relationship

$$HPS = 53.34 (WFT)^{-0.259} \quad (16)$$

was obtained for the dynamic hydroplaning speed HPS (km/hr) as a function of the waterfilm thickness WFT (cm). In U.S. customary units, the equation is

$$HPS = 26.04 (WFT)^{-0.259} \quad (17)$$

for the dynamic hydroplaning speed HPS in mph and the waterfilm thickness WFT in inches. The study noted that considerably more data are required to accurately establish this relationship for waterfilm thicknesses less than 0.25 cm (0.10 in.). However, the critical hydroplaning speed under this condition is much higher, and full dynamic hydroplaning speed is less likely to occur for waterfilms of this depth at legal highway speeds.

LIABILITY FOR HYDROPLANING

All of the previously discussed factors—tire inflation pressure, tread depth and design, pavement texture depth, pavement slope, drainage path length, and rainfall intensity—influence hydroplaning occurrence. But the recognition of environmental conditions creating sufficient water depths on the pavement for the possibility of hydroplaning, and the action of sustaining a reasonable operating speed under those conditions, is the responsibility of the driver.

Loss of control due to high or unsafe speed is the direct cause of most wet-weather accidents. If the driver chooses to ignore high-intensity rainfall and continues to operate at speeds considered high for the existing conditions, the probability of dynamic hydroplaning is increased. With full dynamic hydroplaning, the driver loses control over vehicle steering and braking.

Driver expectations during rainfall must be realistic and reasonable. Operating at posted speed limits greater than 80 km/hr (50 mph) under heavy rainfall places the driver at risk of dynamic hydroplaning. Citations issued by law enforcement personnel in many of these cases charge the driver with operating the vehicle at a "speed unsafe for conditions" or "failure to control speed." Highway engineers must rely on the prudence and reasonable operation of drivers during times of rainfall or when water is on the pavement. Speed should be reduced below 80 km/hr (50 mph) to decrease the probability of full dynamic hydroplaning (*10*). Overt actions or reactions by braking or steering should be carefully controlled when encountering water on the pavement surface, because friction capability is significantly reduced.

Responsibility for proper tire care and maintenance also lies with the driver. Drivers must be relied upon to maintain tire inflation pressures in accordance with the manufacturer's specifications. Although the recommended inflation pressure varies for different types of tires, it is typically at or above 206.85 kPa (30 psi) for most passenger car tires. Tire care and maintenance also imply the driver's responsibility to monitor tire tread wear regularly and to reduce the effects of tread wear on tire performance and safety by properly balancing and rotating the tires at regular intervals. Tire tread depth should be a minimum of 0.159 cm (2/32 in.) to reduce the vehicle's susceptibility to hydroplaning and to obtain optimum wet traction performance (*10*).

Highway engineers have responsibility (liability) for properly designing, constructing, and maintaining the roadway pavement to adequately drain surface water from normally expected rainfalls. This includes the recognition and remediation of pavement defects, failures, or areas prone to the possibility of ponding water. However, as stated previously, under the most desirable methods of design, construction, and maintenance of a roadway for pavement surface drainage, an atypical, high-intensity rainstorm can produce sheet flooding or water ponding such that hydroplaning can occur.

Both transverse and longitudinal areas of water puddling may develop on roadways because of wheel loads or failure of the pavement over time. These "ruts" trap water and are most likely to occur on flexible pavements and be of short length. Studies indicate that hydroplaning can occur in these areas when the length of the rut is 9.144 m (30 ft) or greater. However, with normal cross slopes (≤2.5 percent), rut depths of 0.61 cm (0.24 in.) or less do not significantly contribute to a higher risk of hydroplaning (*11*).

Special attention must be given by highway engineers to areas on roadways prone to ponding of water under high-intensity rainfall rates. Drainage facilities should be emphasized that will rapidly collect and remove water from locations

of flat or sag vertical profile that are susceptible to hydroplaning under heavy rainfall conditions.

Horizontal alignment transition areas with superelevation also may create a "flat spot" in the transverse cross section of a roadway. This is an especially critical point where little or no longitudinal slope exists to drain water away from the traveled way. Highway engineers must anticipate the possibility of ponding water on the pavement in this situation under high-intensity rainfall and introduce drainage adjustments to minimize the probability of hydroplaning.

HYDROPLANING AND ROADWAY TORT LITIGATION

An increasing number of wet weather accidents have resulted in lawsuits with claims of proximate cause being water on the pavement surface inducing loss of control through hydroplaning. The allegations in this litigation may be focused in two areas: encountering sheet flooding or ponded areas of water on the pavement surface and testimony regarding operating speed and loss of control. The following hypothetical legal cases involving hydroplaning and tort liability are presented to illustrate typical allegations versus factual evidence and failure to fulfill duties (negligence) by either the driver or the highway agency.

Case 1

Driver A was proceeding through a right-hand curve on a two-lane, asphalt roadway during a moderate rain shower in daylight. Just before completing the curve, Driver A lost control of the vehicle and crossed the centerline of the roadway, sliding broadside into an opposing vehicle and injuring Driver B. Driver A filed suit against the highway agency, alleging that loss of control was due to hydroplaning, which resulted from a roadway defect.

At the time of the accident, the roadway curve was well marked and signed with an advance curve warning and an advisory speed plate of 64 km/hr (40 mph). Radius of curvature and cross slope (superelevation) were shown to be in compliance for the classification of roadway and posted operating speed. The pavement surface was well traveled, yet shown to have a good coefficient of friction. No record of complaints of comparable accidents at the same curve location were found within a prior 3-year period. Both vehicles were assessed in good mechanical condition, and their tires were in adequate condition and properly inflated.

Driver A testified to a precollision speed below 64 km/hr (40 mph). Damage to both vehicles indicated an impact speed of greater than 80 km/hr (50 mph). The alleged hydroplaning most probably would not have occurred at a speed of 64 km/hr (40 mph) or less at this site under these geometric, pavement, and tire conditions. The broadside skid was also indicative of excessive speed above that posted and critical for the curve alignment.

Case 2

Driver C was traveling on a rural Interstate highway with a posted regulatory speed of 104 km/hr (65 mph) approaching a severe rainstorm. On encountering the rainfall, the vehicle ran off the roadway and struck a tree within the divided median. Driver C sustained injuries in the collision, for which suit was brought against the operating agency alleging hydroplaning to be the cause of loss of vehicle control.

The highway was a four-lane, divided, tangent section at the point of vehicle departure from the roadway. The roadway surface had been recently overlain with asphaltic concrete, providing a high frictional coefficient. Cross slope at the location was measured and found to be in compliance with published criteria.

Driver C testified that he was traveling at 104 km/hr (65 mph) when loss of vehicle control occurred. Other motorists testified to reducing speed to 80 km/hr (50 mph) because of the obvious reduction in visibility and extent of water on the pavement from the rainstorm. Meteorological data indicated the rainfall intensity for the thunderstorm associated with the accident to be near 10.2 cm/hr (4 in./hr) and the cause of flooding damage.

In this case, Driver C may have lost control of the vehicle as a result of hydroplaning upon encountering water on the pavement surface of considerable depth. Driver C may have left the roadway because of poor visibility or may have lost control of the vehicle as a result of inappropriate steering or braking reactions to hydrodynamic forces. However, it is likely that this accident was the direct result of Driver C's failure to recognize and respond to adverse weather conditions. Reasonable and prudent action on the part of Driver C, in the form of a speed reduction, would have likely avoided this accident.

Case 3

Driver D was traveling on a two-lane, asphalt roadway entering a left-hand curve during slight rainfall. Loss of control caused the vehicle to continue in a straight line off an embankment to the outside of the curve. Driver D alleged that water encountered on the roadway caused hydroplaning and the subsequent loss of vehicle control. Suit was brought against the operating agency for negligence in design, construction, and maintenance resulting in a highway defect.

Driver D testified that he was traveling at the posted speed limit of 88 km/hr (55 mph) at the time of the accident. The pavement surface was worn and polished with a marginal, yet adequate, coefficient of friction. The location of the water encountered was determined to be in the superelevation transition from normal, crown cross slope to banked cross slope (superelevation). The transverse grade of an area on the roadway in this transition was measured and determined to be less than 0.05 percent. This "flat" area was compounded by also being at the sag (low) point of a longitudinal vertical grade. Furthermore, evidence indicated an average of five comparable accidents per year for this site for the 3 years before the accident.

For the existing geometric and pavement conditions, it was possible for hydroplaning to have occurred because of water on the roadway for a motorist traveling at the posted speed limit under normally expected rainfall intensities. The path of departure also indicated little or no vehicle control, typical of full dynamic hydroplaning. The agency had a duty and

responsibility to recognize the combination of conditions conducive to poor drainage of the roadway and to remediate those conditions.

CONCLUSIONS

Many roadway tort liability claims are being made with little or no factual basis to substantiate allegations of hydroplaning as a causative factor. The physical phenomena of dynamic hydroplaning can only be possible at a designated minimum speed when water depth on the roadway exceeds the combined surface macrotexture depth and tire tread depth. Other factors of influence, such as tire inflation pressure and tire footprint size and shape, may adjust the calculation of the critical hydroplaning speed.

Highway engineers have responsibility for roadway factors affecting friction capability, such as pavement texture design and depth, and surface drainage, such as cross slope, superelevation transition, longitudinal grade, and length of the transverse drainage path. Engineers must design, construct, and maintain streets and highways in a manner ensuring proper surface drainage to minimize the probability of water accumulation under normal rainfall conditions.

Motorists must also accept responsibility for their driving behavior during periods of rainfall. A reasonable and prudent driver should recognize the greater potential danger of operating a vehicle in a wet roadway environment and reduce vehicle speed to minimize the risk of losing control of the vehicle. For most cases of full dynamic hydroplaning (assuming adequate tire tread and proper tire inflation), the vehicle speed at which hydroplaning is observed would be considered unsafe or not prudent for the amount of water on the roadway.

Judges and juries in cases of roadway tort litigation must determine whether hydroplaning occurred and its relevance as a causative factor in many accidents. In addition, assessment must be made as to responsibility for conditions that result in hydroplaning. These decisions can only be made with factual information about the physical phenomenon of hydroplaning and factors influencing both the roadway and vehicle. It is hoped that this paper has addressed issues relevant to hydroplaning and roadway tort litigation in an informative and helpful manner.

REFERENCES

1. *Fatal Accident Reporting System*. DOT-HS-807-794. NHTSA, U.S. Department of Transportation, Dec. 1991.
2. *Single Variable Accident Tabulations*. Texas Department of Transportation, July 1992.
3. *Road Surface Characteristics: Their Interaction and Their Optimisation*. Organisation for Economic Cooperation and Development, Paris, 1984.
4. J. S. Creswell, D. F. Dunlap, and J. A. Green. *NCHRP Report 176: Studded Tires and Highway Safety: Feasibility of Determining Indirect Benefits*. TRB, National Research Council, Washington, D.C., 1977.
5. R. Pelloli. Road Surface Characteristics and Hydroplaning. In *Transportation Research Record 624*, TRB, National Research Council, Washington, D.C., 1976.
6. W. B. Horne. Status of Runway Slipperiness Research. In *Transportation Research Record 624*, TRB, National Research Council, Washington, D.C., 1976.
7. W. B. Horne and F. Buhlmann. A Method for Rating the Skid Resistance and Micro/Macrotexture Characteristics of Wet Pavements. In *Frictional Interaction of Tire and Pavement, ASTM STP 793* (W. E. Meyer and J. D. Walter, eds.), American Society for Testing and Materials, 1983, pp. 191–218.
8. R. W. Yeager. Tire Hydroplaning: Testing, Analysis, and Design. In *The Physics of Tire Traction: Theory and Experiment* (D. F. Hays and A. L. Browne, eds.), Plenum Press, New York, 1974.
9. A. L. Browne. Mathematical Analysis for Pneumatic Tire Hydroplaning. In *Surface Texture Versus Skidding: Measurements, Frictional Aspects, and Safety Features of Tire-Pavement Interactions, ASTM STP 583*, American Society for Testing and Materials, 1975, pp. 75–94.
10. B. M. Gallaway, F. C. Benson, J. M. Mounce, H. H. Bissell, and M. J. Rosenbaum. Pavement Surface. In *Synthesis of Safety Research Related to Traffic Control and Roadway Elements, Vol. 1*. FHWA-TS-82-232. FHWA, U.S. Department of Transportation, Dec. 1982.
11. G. G. Balmer and B. M. Gallaway. Pavement Design and Controls for Minimizing Automotive Hydroplaning and Increasing Traction. In *Frictional Interaction of Tire and Pavement, ASTM STP 793* (W. E. Meyer and J. D. Walter, eds.), American Society for Testing and Materials, 1983, pp. 167–190.
12. W. B. Horne. Predicting the Minimum Dynamic Hydroplaning Speed for Aircraft, Bus, Truck, and Automobile Tires Rolling on Flooded Pavements. Presented to ASTM Committee E-17 Meeting at the Texas Transportation Institute, College Station, Tex., June 1984.
13. D. L. Ivey. *Truck Tire Hydroplaning—Empirical Confirmation of Horne's Thesis*. Texas Transportation Institute, Texas A&M University System, College Station, Nov. 1984.
14. W. B. Horne, T. J. Yager, and D. L. Ivey. Recent Studies To Investigate Effects of Tire Footprint Aspect Ratio on Dynamic Hydroplaning Speed. *The Tire Pavement Interface, ASTM STP 929* (M. G. Pottinger and T. J. Yager, eds.), American Society for Testing and Materials, Philadelphia, Pa., 1986, pp. 26–46.
15. N. C. Yang. *Design of Functional Pavements*. McGraw-Hill Book Company, New York, 1972.
16. W. B. Horne and U. T. Joyner. Pneumatic Tire Hydroplaning and Some Effects on Vehicle Performance. SAE Paper 970C, Jan. 1965.
17. B. M. Gallaway, D. L. Ivey, G. Hayes, W. B. Ledbetter, R. M. Olson, D. L. Woods, and R. F. Schiller, Jr. *Pavement and Geometric Design Criteria for Minimizing Hydroplaning*. FHWA-RD-79-31. FHWA, U.S. Department of Transportation, Dec. 1979.
18. S. K. Agrawal, W. E. Meyer and J. J. Henry. *Measurement of Hydroplaning Potential*. FHWA-PA-72-6. Pennsylvania Transportation Institute, The Pennsylvania State University, Feb. 1977.
19. R. S. Huebner, J. R. Reed, and J. J. Henry. Criteria for Predicting Hydroplaning Potential. *Journal of Transportation Engineering*, Vol. 112, No. 5, Sept. 1986, pp. 549–553.

Publication of this paper sponsored by Committee on Tort Liability and Risk Management.