TRANSPORTATION RESEARCH RECORD 975

Analysis and Management of Traffic Accident Records

TRB
TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1984

Transportation Research Record 975
Price $7.80
Editor: Scott C. Herman
Compositor: Joan G. Zubal
Layout: Betty L. Hawkins

mode
1 highway transportation

subject areas
51 transportation safety
54 operations and traffic control

Transportation Research Board publications are available by ordering directly from TRB. They may also be obtained on a regular basis through organizational or individual affiliation with TRB; affiliates or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Research Council, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

Printed in the United States of America

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Analysis and management of traffic accident records.

(Transportation research record; 975)
1. Traffic accidents–United States–Congresses. 2. Traffic safety–United States–Congresses. 3. Highway departments–United States–Data processing–Congresses. I. National Research Council (U.S.). Transportation Research Board. II. Series.
HE5614.2.A665 1985 363.1'251 85-299
ISBN 0-309-03757-3 ISSN 0361-1981

Contents

Addresses of Authors

Chirachavala, Thipatai, Accident Analysis Division, Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843; formerly with University of Michigan
Cleveland, Donald E., Department of Civil Engineering, University of Michigan, Ann Arbor, Mich. 48109
Hauer, Ezra, Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada
Kostyniuk, Lidia P., Department of Civil Engineering, University of Michigan, Ann Arbor, Mich. 48109
Larratt-Smith, Mark H., Ontario Ministry of Transportation and Communications, 1201 Wilson Avenue, East Building, Downsview, Ontario M3M 1J8, Canada
Mallory, Bradley L., Pennsylvania Department of Transportation, Transportation and Safety Building, Room 215, Harrisburg, Pa. 17120
Mosher, Clarence W., Office of Alcohol and Highway Safety, New York State Department of Motor Vehicles, Empire State Plaza, Albany, N.Y. 12228
Persaud, Bhagwant N., Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada
Tobin, Douglas K., Bureau of Driver Licensing, Pennsylvania Department of Transportation, Room 104, Transportation and Safety Building, Harrisburg, Pa. 17120

Managing Traffic Records Systems Through Management Information Systems

DOUGLAS K. TOBIN

ABSTRACT

Traditionally, traffic records systems have been defined as management information systems for highway safety programming. In fact, they are also massive operating systems concerned with products, revenue collection, and record maintenance. Highway safety management information is often a secondary concern. Managing these operating systems requires smaller management information systems whose data serve to support, control, define, and analyze the operations of the major traffic records systems. Although these smaller management information systems have not been traditionally thought of within the context of the traffic records system, their success or failure may have consequences for the success or failure of the traffic records system and the entire highway safety programming effort. These small support systems serve to ensure the timely creation and updating of the traffic records systems with quality data. They also serve to ensure performance of basic functions, thereby permitting greater management time and attention to highway safety programming.

Traffic records have traditionally been described as record keeping systems for driver, vehicle, roadway and environment, and accident data. Perhaps even more so, the accident records system has generally been held to be the traffic records system.

In concept, a traffic records system is the management information system (MIS) dedicated to highway safety. In reality, however, traffic records systems, at least the major subsystems of the driver and vehicle, are large extensive processing, production, and revenue collection systems that operate in addition to the highway safety programming uses. In fact, their use as highway safety systems can be secondary, particularly in the case of the vehicle file, to that of revenue collection and other purposes.

The driver and vehicle systems contain millions of records and are extremely dynamic. The ability to quickly create and update these records with quality data is a prime determinant of the value of a state traffic records system.

Pennsylvania has approximately 7.5 million vehicles with a total active vehicle registration file of 13 million records. On any given work day, an average of 44,000 paid transactions are processed against that file. Several thousand additional free transactions are also processed, particularly change of addresses.

The Pennsylvania driver license file contains 7.25 million drivers, while the number of active records on file is approximately 8.8 million. Some 40,000 driver licenses, new and renewed, are produced and 20,000 citations and suspension actions are taken each week. By contrast, Pennsylvania's accident record system creates 130,000 records a year and 500 records daily. Altogether, the active accident record file is about 550,000 records. And even though this system is Pennsylvania's most sophisticated traffic records system, it still has performance standards for its managers that are productivity oriented in an operational sense.

Thus to view traffic records systems solely within the context of providing highway safety data is to really miss the mark by a fair margin.

If the operating systems work and work well, there is a reasonable chance that an integrated traffic records system will pay off on its investment. If the operating system does not work well, budget dollars are going not to an integrated traffic record MIS but instead are going into propping up the operating systems because they interface directly with the motoring public. Because of that, they are never allowed to fail.

To illustrate that point, in 1970 Pennsylvania had no accident record system, even though a computerized accident record system had been operational since 1966. The entire Bureau of Accident Analysis, which operated the accident record system, was shifted lock, stock, and barrel to assist the faltering vehicle registration system that year.

Although the driver, vehicle, and accident record systems are massive and complex, managing the day-to-day operation of these data factories requires smaller MISs, whose data serve to support, control, analyze, and direct. (The roadway information system, which is a vital part of the traffic records system, is excluded from this discussion in order to concentrate only on the driver and vehicle systems.)

Previously, many of these MISs were in reality management reporting systems; that is, how many widgets or transactions were produced. In Pennsylvania the registration of vehicles and the licensing of drivers began in 1906; since then a few documents have been lost in file drawers only to be recovered later and treasured as historical documents. Among these papers are different reports used over the years. They all are of the management reporting variety. Those types of reporting systems remain and in fact were the MISs until 1979. Many published accident statistics are little better than reporting mass arrays of data by frequency and type.

An MIS is no substitute for management; rather it is a tool. It is immaterial whether the MIS is computerized or not. What is material is the relevance of what it measures, the relevance of the data it collects and manipulates. This holds true for the traffic records system as well.

MISs are to report the exception and prove the general rule or standard. For example, the Bureau of Traffic Safety wants a specific population of repeat offenders to be defined. In addition to defining the parameters of that group, the Bureau needs demographics from the files for that population because they are an exception. Likewise, for statistical quality control, the Bureau wants to know when the process is within standard or tolerance and when a process is out of control before it goes beyond a set tolerance.

Whether needing data on recidivism rates for school bus violations, suspensions, or the time to produce a driver's license, the principles for an MIS remain the same:

1. Know what needs to be measured,
2. Set up a collection and reporting system, and
3. Collect the pertinent data.

Oftentimes the order is reversed. The data that are available are reported, not the data that are needed.

Beginning in 1979 at the Pennsylvania Department of Transportation (PennDOT), department-wide reporting began as part of a management-by-objectives program. How much was produced was not a relevant measure of performance because PennDOT had a captive market and it was not relevant to the level of service provided to its customers. Because the Department could not or should not license any more drivers than were eligible, it concentrated on how well the licensing functions were performed. What also became relevant was the cost of operating those systems, particularly in terms of staffing.

To better understand how these subsystem MISs are used to run PennDOT's main operating systems, three examples are discussed.

TURNAROUND TIME REPORTING

Pennsylvania's traffic records systems are fairly unique, at least in one aspect. Although as a general rule accident record systems are centralized, driver and vehicle license issuance throughout the rest of the United States tends to be decentralized. In Pennsylvania, however, all processing and issuance is centralized in Harrisburg. The only transactions done outside of Harrisburg are those for temporary license plates, which are available from dealers, and taking the photograph for driver licenses. All products are produced in Harrisburg and either sent back to Pennsylvania motorists through the mail or by registered messenger services. Thus the time it takes to produce a given product is an important indicator of performance. From a traffic records point of view, it is just as important to be timely in creating and updating records. Over the past 4 years a series of small MISs have been developed and implemented to measure the length of time it takes to complete a given process. These MISs are fairly simple but have a high degree of relevance in terms of system management for performance.

By December 1980 turnaround times were measured for the four major product lines: titles and vehicle registrations, driver licenses, registration renewals, and driver license renewals. At first the systems were extremely simple, with data being extracted from processed source documents. The mail opening receipt date was compared with the automated system process date and processing time was calculated in term of workdays.

In 1980 the title and registration turnaround time system was changed. A sample size of 100 processed documents was selected at random each day to give a 95 percent level of confidence. Processing time was broken down for each discrete action within the whole system. In other words, processing time was calculated for mail opening, another time calculated for sorting, and so on, until the certificate of title had actually been processed.

With this more detailed information, management began to work on the numbers. Numbers that appeared disproportionately high with respect to the activity involved for a particular process in the system were worked on by special management and supervisory teams. Procedures for those processes were either scrapped or revised to permit streamlined processing. A new title and sales tax applications processing system was developed and implemented. One of the important design elements of the new system was an automated MIS for turnaround time. The necessary data storage was created on the vehicle registration file and the requisite transactions were modified to either capture or automatically acquire process date data for each defined processing point.

However, the automated system did not work and has not worked yet. The difficulty was not in capturing the data, but in the logic to determine which transactions and parts of certain transactions were pertinent to determining appropriate turnaround times.

The automated reporting system was quietly turned off. However, because the data were captured by the system, a routine inquiry transaction was used to acquire the data needed by video display terminal (VDT). Processed batches were selected at random as before. However, the sample size was increased to 2,500 because through VDT inquiry, acquisition, and manipulation of data were fairly simple, although still manual.

The turnaround time management information that began with the vehicle title and registration process has been expanded to other areas. It has been the most useful of PennDOT's MISs. The Department measures processing times for processing citations, court records, accident reports, and for processing the client in-take evaluation results for first-offense driving under the influence (DUI) drivers in the court reporting network.

As with other endeavors, time is of the essence when managing the operating systems that comprise an integrated traffic records system. Timeliness of data acquisition and reduction is an important determinant of the quality of highway safety management systems that support highway safety programming.

PRODUCTION PLANNING

The Department faced an annual crisis every spring and summer with the seasonal rise of car sales and the staff taking vacations. Adding to that crisis was the processing of driver license and learner permit applications that uses some of the same resources, such as mail opening and sorting. This activity also has an annual increase in activity at the same time. In the past the solution to this particular recurring crisis was increased staffing. However, with emphasis on staff reduction, the Department knew that it had to plan better both in terms of streamlining the processing system and staffing.

Larry White at the PennDOT Bureau of Motor Vehicles and Licensing and his staff began initial planning in the fall of 1981 by using the previous year's volumes as a model for the anticipated work flow through the various working areas. The initial application was used successfully to schedule annual leave in affected work areas.

In 1982, after acquiring an Apple III microcomputer, work began on a computerized production forecasting system for the work areas of the title and registration and driver license application systems.

Robert Baron of the Bureau of Motor Vehicles and Licensing developed the forecasting model using Visicalc. The model was designed in part to help clear external variables such as environmental changes that are contingent on political and economic conditions. A large data base was developed from historical and current data from the affected areas.

The most unpredictable factor is the external variable, volume. Volume is forecasted by taking a weighted moving average of demand. Each month the average of the last 12 months is recomputed by adding the latest month's figures and recalculating the average. The next refinement is to give greater weight to the latest information and to calculate a weighted moving average. By using a smoothing constant, the weight given to a period is reduced by a fixed proportion each time the average is recomputed.

The value of the smoothing constant usually lies in a range between 0.01 and 0.09. A low value makes the forecast slow to respond to changes in demand, and a high value makes it react quickly. In other words, if more value is necessary from the most recent data, a smoothing constant of 0.02 or greater is needed. The value used is estimated and is based on forecast error and the known accuracy of the historical data.

This statistical model requires 2 or more years of data for accuracy, with the optimum being 5 years. The average of the last 12 months is recomputed monthly by adding the latest data and then recalculating the average:

New forecast = (Forecast for current period) + [(Current period demand) - (Forecast for current period)] x (Smoothing constant).

Seasonal trends that affect processing are identified to reveal any major differences between peak monthly demand and average demand throughout the year. The forecast is adjusted for seasonal trends by the ratio of each month's demand to the annual average (Figure 1):

[(Average month) ÷ (Adjusted forecast)] x Forecast = Seasonally adjusted forecast.

Internal variables incorporated into the production control reports include sick leave, annual leave, productivity rate per man-hour, employee hours available, and employees available. Sick leave is computed by using a moving average, with the new average computed monthly. Because annual leave percentages fluctuate, the production report contains leave figures at 0, 4, 9, and 14 percent, and forecasts staffing needs at each one of these rates. Productivity rates were taken from production reports of the various units. These rates are constantly checked and updated because of the phenomenon of the learning curve, new procedures, calculating techniques, new employees, and various other factors (Figure 2).

As a result of the production forecast, line managers and supervisors have a fairly accurate idea of the volume of work that can be expected in the near future. This has enabled PennDOT's management and supervisory team to take staff from areas of low demand for use in areas of high demand. It has also been useful for scheduling overtime far enough in advance to give employees time to make arrangements, thereby increasing the number of employees turning out for overtime. And the production forecast is still used for its original purpose: determining when annual and personal leave days may be taken by employees.

Of course the true test of the effectiveness of a system is its use by management and the results obtained from that use. By that measure, this MIS has been an unqualified success. The forecasting accuracy has been in excess of 90 percent, given that only 2.5 years of historical data are available and that in 1983, for the first time in 3 years, there was a significant increase in the volume of work because of increased automobile sales.

The real proof, however, is that for the first time title and registration turnaround time has been less than 8 days during the spring and summer months. In June 1982 this turnaround time was 15.4 days; in June 1983 the turnaround time was 7.6 days,

PRODUCTION FORECASTING REPORT
DRIVER'S LICENSE SECTION (EXAMINING)

MONTHLY FORECAST
FOR NEXT 12 MONTHS

MONTH	1980	1981	1982	1983	1984	Total	Average		1983	1984
Jan		60803	66862	59376	71659	258700	64675	JAN		64227
Feb		67333	82981	67179	79652	297145	74286	FEB		73772
Mar		90375	104445	86989	79422	361231	90308	MAR		89683
Apr		95673	94888	74212	74796	339569	84892	APR		84305
May		86454	98318	80353		265125	88375	MAY		87763
Jun		96868	102958	92184		292010	97337	JUN	96663	
Jul	84765	96852	91142	82886		355645	88911	JUL	88296	
Aug	99107	81034	104877	99107		384125	96031	AUG	95367	
Sep	75012	89757	86817	82954		334540	83635	SEP	83056	
Oct	89289	60361	82983	85795		318428	79607	OCT	79056	
Nov	64001	74723	73604	82384		294712	73678	NOV	73168	
Dec	64444	76100	59376	68030		267950	66988	DEC	66524	
Total	476618	976333	1049251	961449	305529	3769180				
Average	79436	81361	87438	80121	76382	314098				

AVERAGE YEAR*	983264
CURRENT DEMAND*	979222
CURRENT FORECAST*	973697
NEW FORECAST	976460

1. This section contains volumes received information submitted by the units, totals yearly volumes to the current month (total) and also gives average monthly volume for each year (average).

2. The report also contains the total for all homogeneous months, i.e., January, 1982; January, 1983. It also takes the average of the homogeneous months so that the forecast can be seasonally adjusted.

3. Predicted volumes for the next 12 months are given.

4. The average yearly volume, current demand (the last 12 months), and the current and new forecasts are represented. The new forecast represents the volume predicted for the next 12 months.

FIGURE 1 Production forecasting report, driver's license section.

JUNE
PRODUCTION FORECAST

DISCRETIONARY HOURS

PREDICTED-VOLUME	96663	HOURS AVAILBLE AT	ADDITIONAL HOURS NEEDED (NEEDED HOURS-AVAILABLE HOURS)
PRODUCTIVITY RATE- PER MAN HOUR*	44	.05%SICK LEAVE.04% ANNUAL----- 2140	57
HOURS NEEDED	2197	.05%SICK LEAVE.09% ANNUAL----- 2023	174
DAYS IN MONTH?	21	.05%SICK LEAVE.14% ANNUAL----- 1905	292
AVG. NO. OF EMPLOYEES AVAILABLE?	16	.00%SICK LEAVE.00% ANNUAL----- 2352	-155
EMPLOYEE HOURS AVAILABLE	2352	.05%SICK LEAVE.00% ANNUAL----- 2234	-38
CURRENT SICK LEAVE USAGE?	.05		

(-)MINUS INDICATES SURPLUS HOURS

This report provides the managers with a quick analysis of what is taking place within their units.

1. Predicted volume - forecasted volume for month indicated.
2. Productivity rate per man hour - volume of work completed in one hour by an individual.
3. Hours needed - this indicates the amount of time necessary to process the predicted volume.
4. Days in month -'number of working days in month indicated.
5. Average number of employees - number of producing employees within a unit available.
6. Employee hours available - number of employee hours available (based on 7 hours).
7. Current sick leave usage - forecasted sick leave for division.
8. Hours available at current sick leave - sick leave usage minus number of hours available.
9. Discretionary hours -
   - Hours available at - includes hours available at current sick leave and annual leave usage.
   - Additional hours needed - these figures indicate additional hours needed or surplus hours available.
10. Unit capability - this figure indicates what a unit can potentially produce for the month indicated at current productivity rates and staff complement.

FIGURE 2 Productivity rates.

and that was done with less staff in 1983 than in 1982. The numbers are just as dramatic for driver licensing applications.

STATISTICAL QUALITY CONTROL

One of the most important areas of concern for the Department is product quality. In the past traditional quality control methods have been used in an attempt to produce better quality products. There was a proofreading staff that read all titles before they left the Department. There was, and is, a dedicated staff to review accident records processing. The proofreading staff represented 100 percent inspection, and like all 100 percent inspection schemes, it did not work well, and by 1979 had been dropped entirely.

But because there was no longer a quality control activity for title processing, this did not mean that the concern disappeared. In 1980, using the accident record review as a model, PennDOT began quality review by sampling techniques in critical areas of the title processing process. In 1982 a quality control staff was assembled and the error potential of the driver and vehicle operating systems was assessed.

The title and registration system was by far the most error prone. It was not startling news. The automated title and registration system is not particularly quality conscious or quality supportive. In September 1982 a capability study by Constance Whitmarsh-Tomko of the Bureau of Motor Vehicles and Licensing had been completed for all phases of the vehicle title and registration system. Control charts or p-charts were developed to show fraction rejected (Figure 3) for applications processed.

Statistical quality control techniques were chosen because of the failure of the traditional quality control model and a conviction that this was the only technique that had any hope of substantially increasing product quality. Although the Irving Trust Company of New York City has used the technique since 1975 with great success, use of statistical quality control for normal clerical operations is still rare.

To support the quality control plan and to provide management information on the outgoing quality level of the Department's products, a standardized error reporting system was developed by William Hutchinson and Ms. Whitmarsh-Tomko. Before development of this system, error reporting had been done on a unit-by-unit basis without any standardized procedure or idea of the necessary sample size. There was no qualitative measure to the particular error; any type of error rendered the entire transaction in error.

In developing the standardized error reporting system, an error reporting study was done that developed a data base of all total possible errors that could occur in the vehicle title and registration system. Errors were broken down by section and were further divided into type, impact of severity, and frequency, and a procedure for standardizing errors was then developed.

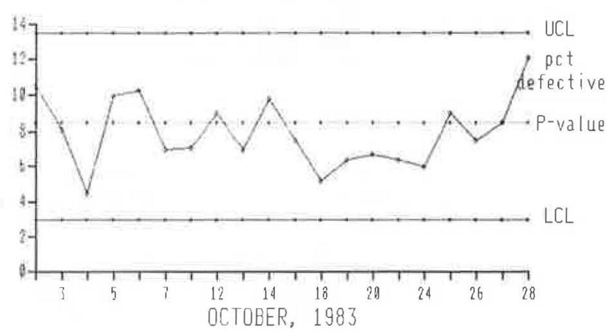

FIGURE 3 Data entry p-chart.

OPERATOR	ERROR/SAMPLE	PERCENTAGE	Q-SCORE
SUP.	0/0	0	0
122	3/140	2.14	70
102	6/140	4.29	61
103	1/135	.74	76
104	6/135	4.44	61
RET	0/30	0	74
106	3/130	2.31	70
121	4/135	2.96	67
113	0/25	0	74
RET	0/35	0	74
115	1/140	.71	76
118	0/140	0	79
119	0/20	0	73
120	5/125	4	63

TOTAL ERRORS = 29
TOTAL SAMPLE = 1330

PROCESS AVERAGE: 2.18
PROCESS CAPABILITY: 2.18

FIGURE 4 Section report giving errors and quality score by operator.

Four error-type categories were established: typing, document mishandling, judgment, and machine. All errors were assigned to one of these four types. The error types were divided by impact. Impacts were either public or system. An error affects the public if it produces a product that will not satisfy a customer's expectations. All other errors are system impact errors; that is, they delay the flow of documents and create additional work.

Finally, errors were then grouped according to their severity. Each was placed in an impact classification, which is different than the impact type in that it gives the degree of impact the error has on the system. Three classifications were used: critical, major, and minor. Critical defects will affect usability, major defects might affect usability, and minor defects will not affect usability of the output piece (1). To distinguish between errors and different impact classification, each was assigned a numerical weight, where critical = 10, major = 5, and minor = 1. Although any weight can be assigned to the classifications, these numbers give sufficient distinction among classifications and are generally considered satisfactory (1).

Once grouped by error type, impact type, and impact classification, it was necessary to develop a means for standardizing errors. All sections cannot commit the same number of possible errors; therefore, a procedure was developed to put all sections on an equal error scale. The net result was an error conversion number assigned to each error. The error conversion number will be higher per error for sections capable of committing few errors and lower per error for those capable of committing many errors. The importance of the error conversion number lies in its balancing out the probability that the section capable of committing many errors will commit more than one capable of committing few errors.

EXAMINING DIVISION
WEEK ENDED: 84/08/03

WORK COMPLETED: 47273
SAMPLE SIZE: 1180
ERRORED APPS.: 36
AVG. ERRORS PER APP.: 1.14

DEGREE OF IMPACT	ERROR DESCRIPTION	WEIGHTED FACTOR	ERROR FREQUENCY	WEIGHT ADJUSTED ERROR
VTR.TAXI	ERRORED APPS: 5 AVG. ERRORS PER APP: 1.2			
CRITICAL	1. incorrect make code	0.313	0	0.000
	2. incorrect registration expiration	0.313	1	0.313
	3. incorrect title code	0.313	0	0.000
	4. failure to reject for additional fee	0.313	1	0.313
	5. failure to reject for proof of ownership	0.313	0	0.000
	6. incorrect tag type code	0.313	0	0.000
	7. incorrect completion of tax print	0.313	0	0.000
	8. incorrect output indicator	0.313	0	0.000
	9. unauthorized missing application	0.313	0	0.000
	10. app inserted in wrong batch	0.313	0	0.000
	11. OTHER	0.313	0	0.000
	SUBTOTAL		2	
MAJOR	12. incorrect totaling of fee	0.156	0	0.000
	13. incorrect rejection code	0.156	0	0.000
	14. unnecessary rejection	0.156	0	0.000
	15. unnecessary rejection, should be special handl	0.156	0	0.000
	16. failure to code for special handling	0.156	0	0.000
	17. fail to rej /other than fee or prf of ownershp	0.156	1	0.156
	18. OTHER	0.156	0	0.000
	SUBTOTAL		1	
MINOR	19. absence of required err cde on rej/not checkin	0.031	0	0.000
	20. rejection and special handling marked	0.031	0	0.000
	21. recorded incorrect fee	0.031	0	0.000
	22. incorrect sales tax computation	0.031	0	0.000
	23. recorded fee not required	0.031	0	0.000
	24. incorrect transaction code	0.031	1	0.031
	25. failure to record required fee	0.031	0	0.000
	26. failure to stamp ID number	0.031	2	0.063
	27. incorrect special handling error code	0.031	0	0.000
	28. failure to code exemption reason number	0.031	0	0.000
	29. failure to code uncommon make	0.031	0	0.000
	30. recorded fee in incorrect place	0.031	0	0.000
	31. wrong proc fee reason cde/fail to cde reason	0.031	0	0.000
	32. OTHER	0.031	0	0.000
	SUBTOTAL		3	
	SECTION TOTAL		6	

FIGURE 5 Section report by error type, frequency, and impact.

The error conversion number is used along with the impact classification weight and a specific error frequency (error frequency data being gathered through quality assurance sampling) to calculate a weight adjusted error. The computation (Error frequency x Error conversion x Impact classification weight) is simple multiplication and gives an adjusted error rate that can be compared across division and section lines. The higher the magnitude of the weight adjusted error, the more serious are the implications of commiting that error.

This information is presented in a hierarchy of four formatted reports based on organizational level, section, division, summary, and management summary error reports. Each successful report represents the data in a manner useful to different levels and applications of management (see Figures 4-7).

This system began in September 1983 after intensive training of managers, supervisors, and lead workers by the quality control staff. The system has been on-line for only a short time, and the results are not readily available. Information from this system can now be used by managers and supervisors to gauge the quality performance of the title and registration system and for development of targeted remedial training of staff and agents. In addition, the data are being used to program system enhancements for quality improvements.

PennDOT believes that this system holds a lot of promise. Its success with the most error-prone system will mean extension of statistical quality control to other traffic records systems. In driver licensing, PennDOT is just beginning to study the possible use of statistical quality control method-

	84/06/29		84/07/06		84/07/13		84/07/20		84/07/27		84/08/03		TOTALS	
	ERRORED APP.	(%)	ERRORED APP.	(%)	ERRORED APP.	(%)	ERRORED APP.	(%)	ERRORED APP.	(%)	ERRORED APP.	(%)	ERRORED APP.	(%)
EXAMINING	40	2.95%	22	2.65%	39	3.17%	23	1.99%	39	2.93%	36	3.05%	199	2.81%
VTR.TAX1	7	2.59%	3	1.67%	4	1.78%	1	0.51%	7	2.86%	5	2.33%	27	2.03%
VTR.TAX2	14	5.83%	6	3.75%	9	3.75%	11	4.78%	7	3.33%	9	4.00%	56	4.29%
VTR.TAX3	10	3.85%	7	5.00%	9	3.91%	4	1.54%	17	4.26%	12	4.53%	59	3.80%
VTR.TAX4	3	1.18%	3	2.00%	8	3.81%	2	1.03%	3	1.54%	3	1.67%	22	1.86%
VTR.TAX5	6	1.82%	3	1.50%	9	2.77%	5	1.82%	5	1.79%	7	2.37%	35	2.05%
DATA ENTRY	75	5.57%	70	6.97%	71	4.88%	99	6.83%	83	5.57%	83	5.55%	481	5.84%
VTR.TAX1	16	6.81%	16	8.42%	10	3.92%	14	5.60%	15	5.56%	14	5.28%	85	5.80%
VTR.TAX2	8	3.17%	8	5.52%	10	3.92%	17	7.23%	7	2.86%	10	3.85%	60	4.31%
VTR.TAX3	4	1.78%	15	8.11%	14	6.09%	13	6.34%	9	3.91%	8	3.64%	63	4.86%
VTR.TAX4	21	6.67%	15	6.98%	6	2.11%	16	5.52%	15	5.56%	14	4.83%	87	5.23%
VTR.TAX5	20	8.70%	6	4.00%	15	6.00%	12	5.00%	14	5.49%	10	4.88%	77	5.79%
VTR.TAX6	6	6.67%	10	8.33%	16	8.89%	27	11.74%	23	10.45%	27	10.59%	109	9.95%
TOTAL ERRORED APPS		115		92		110		122		122		119		680
SAMPLE QUALITY RATE		95.74%		94.99%		95.90%		95.32%		95.67%		95.55%		95.56%
PREDICTED QUALITY RATE		94.98% - 96.5%		93.99% - 95.99%		95.15% - 96.65%		94.51% - 96.13%		94.92% - 96.42%		94.77% - 96.33%		95.23% - 95.89%
PREDICTED ERRORED APPS		3130 - 4490		2308 - 3460		3083 - 4463		3103 - 4402		3168 - 4496		3148 - 4485		20282 - 23539
APPLICATIONS PROCESSED		89439		57566		92030		80178		88505		85764		493482

FIGURE 6 Weekly summary report for full quality control reporting.

	CRITICAL		MAJOR		MINOR		TOTALS		
	ERROR FREQUENCY	WT.ADJ. ERROR	ERROR FREQUENCY	WT.ADJ. ERROR	ERROR FREQUENCY	WT.ADJ. ERROR	ERROR FREQUENCY	WT.ADJ. ERROR	ERRORED APPS.
EXAMINING	72	22.500	84	13.125	58	1.813	214	37.438	199
VTR.TAX1	16	5.000	4	0.625	9	0.281	29	5.906	27
VTR.TAX2	24	7.500	21	3.281	14	0.438	59	11.219	56
VTR.TAX3	18	5.625	26	4.063	18	0.563	62	10.250	59
VTR.TAX4	9	2.813	6	0.938	8	0.250	23	4.000	22
VTR.TAX5	5	1.563	27	4.219	9	0.281	41	6.063	35
DATA ENTRY	256	117.315	177	45.903	70	4.405	503	130.185	481
VTR.TAX1	45	16.667	31	5.741	11	0.407	87	22.815	85
VTR.TAX2	35	12.963	21	3.889	8	0.296	64	17.148	60
VTR.TAX3	33	12.222	13	2.407	17	0.630	63	15.259	63
VTR.TAX4	40	14.815	27	5.000	24	0.889	91	20.704	87
VTR.TAX5	49	18.148	29	5.370	1	0.037	79	23.556	77
VTR.TAX6	54	20.000	56	10.370	9	0.333	119	30.704	109

TOTAL WORK PROCESSED: 493482

TOTAL SAMPLE SIZE: 15321

TOTAL ERRORED APPLICATIONS: 680

ERROR RATE: 4.44%

FIGURE 7 Summary report.

ologies to improve the quality of Pennsylvania drivers.

These MISs are just a sample of the smaller MIS used to operate the driver and vehicle record systems. Although these are specifically pertinent to Pennsylvania, the concept is applicable to any jurisdiction.

These systems have enabled Department management to focus on pertinent and highly relevant problem areas. They have helped measure and guide solutions. The greatest gift, however, has been time. With the operating systems working fairly efficiently, less time is spent by traffic records managers solving crises. Oftentimes these small MISs highlight problems well in advance of the crises stage while their solution is still fairly simple and quick. In addition, the data from these systems often point to the solution.

With management time available, instead of consumed by endless rounds of "firefighting," managers can structure, plan, and nurture that other part of traffic records--the safety MISs.

REFERENCE

1. D.H. Besterfield. Quality Control: A Practical Approach. Prentice-Hall, Englewood Cliffs, N.J., 1979.

Pennsylvania Driving Under the Influence Extra Enforcement Grants: How Traffic Records Can Assist a Highway Safety Program

BRADLEY L. MALLORY

ABSTRACT

From 1982 through 1983 the Pennsylvania Department of Transportation has funded 25 driving under the influence (DUI) extra enforcement grants. These grants consisted of patrol units of one or two officers dedicated solely to enforcing DUI laws. The hours of operation of the units, generally 10:00 p.m. to 4:00 a.m. on weekends, were suggested by data contained in Pennsylvania's accident records system. The 25 counties that received grants were identified from data contained in Pennsylvania's accident records system. A highway safety planning tool called the municipal accident priority system was used to generate a list of Pennsylvania's 67 counties in descending order of their alcohol-related accident problem. Originally, proposals were solicited from the top 20 problem counties. Thirteen counties responded and received grants. In the second phase proposals were solicited from the second group of 20 problem counties. Twelve of these counties received grants. The extra enforcement grants have resulted in increases in total DUI enforcement levels, ranging as high as 410.53 percent. The cost per arrest under the grants ranged from $220.28 to $613.51 during the hours specified. Preliminary accident statistics suggest that accident activity has decreased more in the municipalities with DUI extra enforcement grants than in those municipalities that did not have grants. This is encouraging, but further research will be necessary to more exactly determine the contribution of increased DUI extra enforcement to decreased accident activity.

There is general agreement in the highway safety community that an increased level of enforcement is the single most effective countermeasure to reduce the number of alcohol- and other drug-related accidents. The theory is that increased enforcement deters people from driving drunk by making them believe that they will be caught if they do.

The nationwide average level of driving under the influence (DUI) enforcement was approximately 1.8 arrests per officer in 1982 according to NHTSA. Some highway safety experts have suggested that an average of at least 2.0 arrests per officer would be necessary to have any meaningful impact. However, there is little, if any, empirical evidence to support this proposition.

Enforcement rates vary greatly from state to state and even within states. Pennsylvania has traditionally been at the low end of the spectrum. Figure 1 shows the DUI enforcement in Pennsylvania from 1978 through 1982. The level remained fairly consistent through 1980 at or below 0.8 arrests per full-time officer.

However, Pennsylvania's level of DUI enforcement increased 37 percent (from 0.84 to 1.15 arrests per officer) from 1980 through 1982. There are two main reasons for this increase. First, police officers and management were responsive to heightened public interest in the DUI problem. In response to this heightened interest, Governor Dick Thornburgh ap-

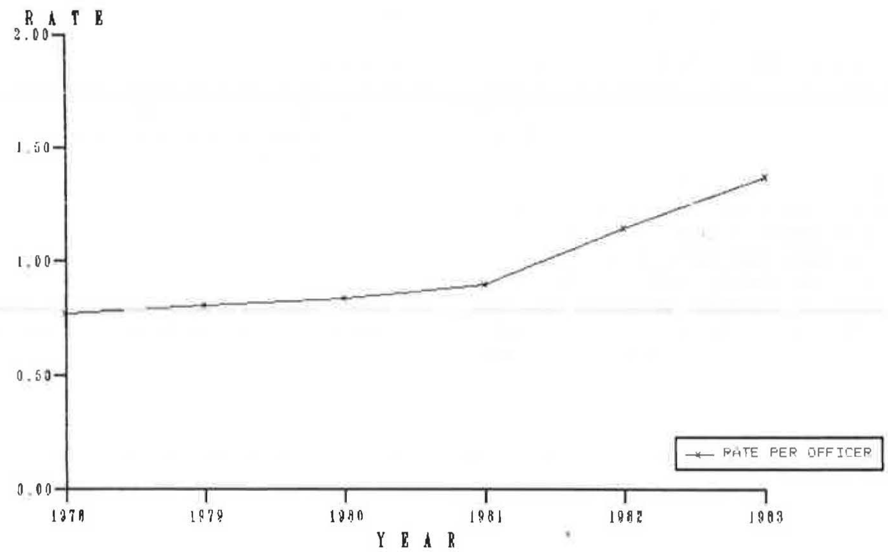

FIGURE 1 DUI arrest rate for Pennsylvania, 1978-1983.

pointed a Task Force on Driving Under the Influence of Alcohol and Other Controlled Substances in 1981. Two local police chiefs and the Pennsylvania State Police Commissioner were Task Force members.

The second reason for the state's improved arrest rates is a program specifically designed to increase DUI enforcement. In its final report, the Governor's Task Force recommended that DUI enforcement activities be expanded. The Task Force noted that data from Pennsylvania's accident records system indicated that most drunk driving and DUI-related accidents occurred on weekends, during the late night or early morning hours (see Figure 2). Usually, traffic patrols are assigned to peak traffic hours rather than those identified as peak DUI hours. Existing late-night patrols had heavy general crime prevention duties that prevented them from focusing on DUI enforcement.

EXTRA ENFORCEMENT GRANTS

In Pennsylvania the highway safety program is administered by the Pennsylvania Department of Transportation (PennDOT). During 1982 the Department decided to create a series of DUI extra enforcement grants by using federal highway safety funds. The grants would pay the salaries of special DUI teams. These teams would consist of one- or two-officer patrol units and would operate during peak DUI hours. No grant monies would be used for clerical, administrative, or equipment costs.

There are 67 counties and more than 2,500 municipalities in Pennsylvania. Few counties have countywide traffic law enforcement units of the type found in many states. Most local traffic law enforcement, including DUI, is conducted by municipal police. It would have been unwieldy to solicit grant proposals from more than 2,500 municipalities, process the information, select the recipients, and administer a myriad of grants. As a result, countywide grants were selected for ease of administration and evaluation. The Department required a county project director for each grant. This project director would solicit municipal participation in their county and

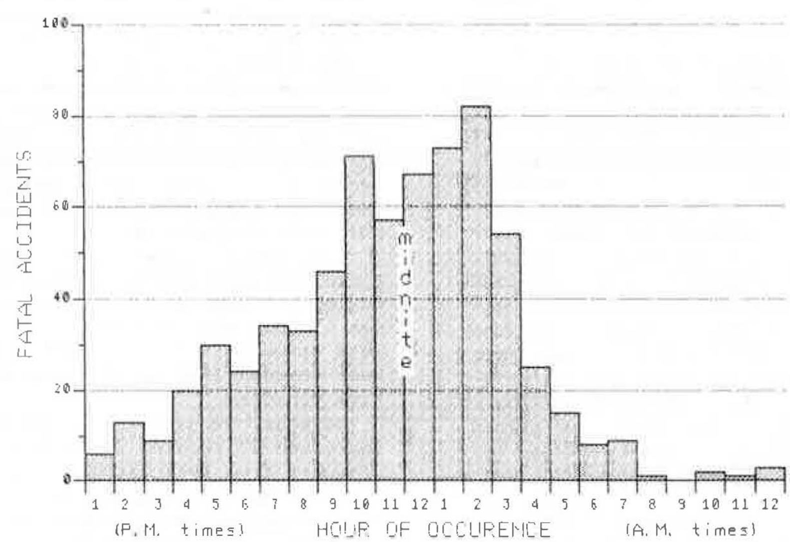

FIGURE 2 Alcohol-related fatal accidents by hour of day, 1982.

coordinate the activities of these subgrantees. Lorraine Novak of the Bureau of Safety Programming and Analysis of PennDOT provided overall program coordination and direction.

PennDOT solicited grant proposals from 20 counties. These 20 counties were selected by ranking all 67 counties by the severity of their DUI-related accident problem. The ranking was developed from data contained in Pennsylvania's accident records system.

ACCIDENT RECORDS SYSTEM

The accident records system in Pennsylvania is a computerized file that contains data from all reportable traffic accidents occurring in Pennsylvania. A reportable accident is defined by law as an accident that results in a fatality (within 90 days), an injury to any person involved, or an accident that results in damage to any vehicle to the extent that it must be towed away.

Police officers are required by law to investigate all reportable accidents and submit standardized accident report forms to PennDOT. Data from these reports and other information sources are entered into the accident records system. Other sources include driver licensing files, vehicle registration files, coroner's and medical examiner's reports, the Pennsylvania roadway information system, municipal maps, and straight-line diagram maps. Analysts enter the data via a terminal directly into the computerized file. The information is automatically edited on entry for range, verification, and consistency.

The accident records system contains descriptions in a standard format of each accident reported. This format contains almost 100 data elements that characterize various attributes of the accident, including vehicles and persons involved, weather and highway conditions, and location information. The format provides sufficient detail to identify hazardous locations and to plan necessary modifications. It also generates the statistics necessary to plan safety programs.

One of the accident records system tools used in planning safety programs is the municipal accident priority system (MAPS). This system ranks municipalities based on aggregated 3-year accident history. MAPS calculates mileage, population, and accident severity rates that are then compared with either countywide or statewide average rates and ratios. The ratios are combined to reach a final point assessment. This assessment is used to rank each political subdivision, either within its county or within the state. Counties may also be ranked according to their relative position within the state. A variety of rankings can be obtained by varying population or road-mileage parameters or by inputting only certain types of accidents.

The data contained in an accident priority listing by county are as follows:

Accident Priority Listings by County for 19__ to 19__

COUNTY	TOTAL ACC.	TOTAL MIL	TOTAL POP.	ACC/MIL	ACC/POP
--	------	----,--	-------	--,--	-----,--

RAT/MIL	RAT/POP	RAT/SEV	POINTS
--,--	--,--	--,--	-----,--

RAT/MIL is defined as the ratio of intensity of accidents per mile of highway in a county to the intensity of the state base (to two decimal places):

$$RAT/MIL = (ACC/MIL)/(BASE\ ACC/MIL) \quad (1)$$

RAT/POP is defined as the ratio of the rate of accidents per 100,000 population in a county to the rate of accidents per 100,000 of the state base data (to two decimal places):

$$RAT/POP = (ACC/POP)/(BASE\ ACC/POP) \quad (2)$$

RAT/SEV is defined as the ratio of the average probable accident severity of a county to the average probable accident severity of the state base data (to two decimal places). The average severity is determined by applying a calculated or relative severity point rating for each accident to the total number of accidents of that description. The totals for each accident description are summed and divided by the total number of accidents to obtain the average severity:

$$RAT/SEV = (AVG\ SEV)/(BASE\ AVG\ SEV) \quad (3)$$

POINTS are the final basis for ranking. They are determined by adding the ratio of mileage rates to the ratio of population rates and multiplying by the ratio of average severity:

$$POINTS = (RAT/MIL + RAT/POP) \times RAT/SEV \quad (4)$$

PROGRAM

The application of MAPS to the Pennsylvania alcohol-related accident records resulted in a list of Pennsylvania's 67 counties in descending order of their alcohol accident problem as indicated by the POINTS. Invitations to submit proposals for DUI extra enforcement grants were sent to the top 20 accident problem counties. These invitations were mailed to county drug and alcohol, probation, or District Attorney's offices depending on the structure of the DUI program in the county. Detailed guidelines for the proposals specified that they should contain a brief problem statement, quantitative goals and objectives (e.g., increase DUI arrests by 50 percent), program description, administrative detail, data-collection techniques, and a budget.

Thirteen counties submitted proposals. After review and some supplemental information, all 13 counties received DUI extra enforcement grants. The 13 original extra enforcement grants were set at 6 months duration and ran from September 1982 through February 1983. In March 1983 all 13 grants were extended another 6 months through August 1983. In addition, invitations for proposals had been extended to the next 20 counties in terms of accident problems or identified by MAPS. Twelve of these counties also began 6-month DUI extra enforcement grants in March 1983. The 25 counties selected and the amounts of their grants are given in Table 1.

TABLE 1 Extra Enforcement Grants and Amounts

County	Amount ($)	County	Amount ($)
Allegheny	84,000	Delaware	49,920
Armstrong		Erie	49,920
Beaver	50,835	Fayette	25,925
Berks	25,920	Franklin/Fulton	27,000
Blair	35,250	Lancaster	24,576
Bucks	33,280	Lebanon	56,768
Butler	29,636	Lycoming	
Carbon	24,336	McKean	54,536
Chester	20,827	Schuylkill	29,280
Columbia	17,880	Warren	24,960
Crawford	25,896	Wyoming	15,552
Dauphin	62,400	York	66,205

TABLE 2 Percentage Changes in DUI Arrests for 1981-1982 and the Average Cost per Arrest

County	Change in DUI Arrests 1981-1982 (%)	Avg Cost per Arrest ($)
Berks	+36.18	256.93
Blair	+9.17	373.87
Carbon	+137.50	221.94
Columbia	+1.22	291.32
Crawford	+57.89	613.51
Franklin	+3.82	346.54
Fulton	+141.67	220.28
Lebanon	+64.18	275.95
McKean	+410.53	308.67
Schuylkill	+34.10	351.37
Warren	+50	234.63
Wyoming	+153.8	391.40
York	+90.28	221.55

The results of the DUI extra enforcement grants have been quite satisfying. The total number of DUI arrests in participating counties has increased as much as 410 percent. The data in Table 2 give the percentage increases in each of the original 13 counties from 1981 to 1982. An average cost per arrest during the target hours (for a 6-month time frame) under the grant is also specified.

The 13 DUI extra enforcement grants operative in the last 3 months of 1982 contributed to the improved statewide DUI arrest picture for that year. Figure 3 shows that DUI arrests finally began to improve significantly in 1981 and particularly in 1982, despite declining police personnel.

The accident experience of the participating counties and municipalities has also been positive. The data in Table 3 give the percentage differences for types of accidents (e.g., alcohol, nighttime, fatal) between the first 4 months of 1982 and 1983, according to statewide accident records figures and data from municipalities that had DUI extra enforcement grants during that time period.

CONCLUSIONS

It is too early to make statistically significant statements about the impact of the grants on the frequency of alcohol-related accidents. The alcohol-related accident trends for the latter part of 1982 and early 1983 are encouraging. Corresponding data for late-night fatal accidents indicate the first downturn in recent years. These reductions cannot be solely attributed to increased enforcement. Many other factors, including the formation of the Governor's Task Force and the publicity surrounding its deliberations, can influence the accident trends, as can the enactment of a new DUI law.

A target figure of two arrests per officer per year has been used in the past as a desirable goal. There appears to be no factual basis for this number in terms of producing a desired reaction in the number of drinking drivers. Another rate commonly mea-

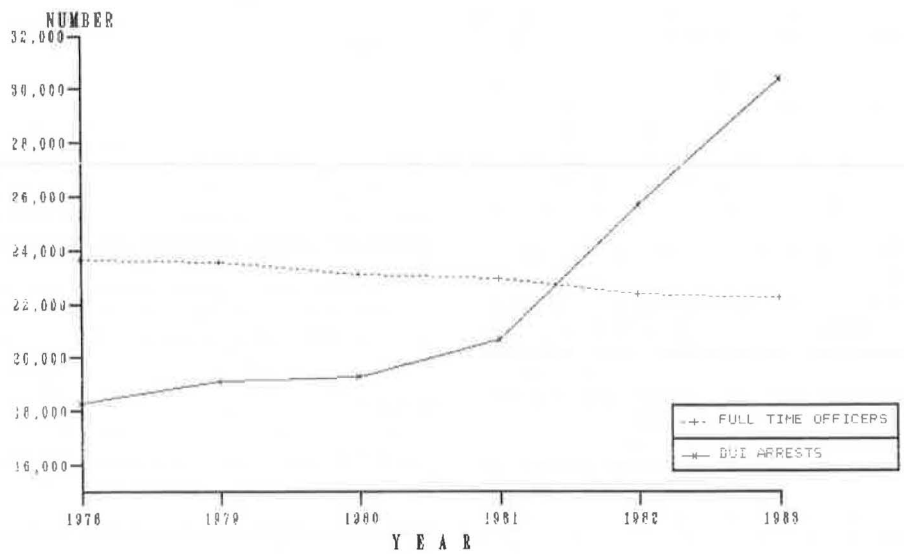

FIGURE 3 DUI arrests in Pennsylvania, 1978-1983.

TABLE 3 Percentage Change Between First 4 Months of 1982 and 1983 for Use in Evaluating DUI Extra Enforcement Grants

	Change in Municipalities with DUI Extra Enforcement Grants Within the County (%)	Change in Municipalities Without DUI Extra Enforcement Grants Within the County (%)	Change Statewide (%)
Fatal accidents	-23.33	0	-1.59
Injury accidents	-6.91	-2.7	-3.04
Property-damage-only accidents	-15.55	-10.65	-11.09
Total accidents	-10.55	-5.72	-6.12
Total fatalities	-24.24	-3.06	-4.48
Total injuries	-5.01	-3.31	-3.44
Alcohol accidents	-8.93	+2.74	+1.54
Nighttime accidents	-17.12	-9.58	-10.16

sured is the number of arrests per 1,000 licensed drivers. Again there appears to be no identifiable rate at which a desired reaction in drinking drivers will occur.

There is a significant lack of research on enforcement rates versus the reaction of drivers. It may well be that there is indeed no ideal enforcement level and that rates of change in enforcement (or perceptions of change) may be the only factor that influences drivers. There is some basis for this hypothesis, in that early peaks of reaction are commonly seen in increased enforcement efforts with a subsequent rapid tailing off, even when higher enforcement levels are maintained. The well-known English experience is an excellent example.

If research were to find that rates of change rather than actual levels of enforcement were producing the desired reactions in the driving public, this would have a significant impact on future enforcement strategies. Lacking research on this topic, researchers must continue to strive for an ideal enforcement level that attempts to balance reactions with resources.

It is certain that increased enforcement must be accompanied by significant efforts. As stated at the outset of this paper, drivers must have a perception of taking a significant risk if any enforcement level or increased enforcement activity is to be effective. Even if DUI arrests were increased 1,000 percent, if drivers are not made aware of this fact, administrators should not expect much in the way of lasting impact on accidents or the frequency of drunk driving. Grants should be awarded with fanfare. Media cooperation in publicizing not only the grant but its results should be obtained.

Publicity and increased enforcement must work together, as neither can stand alone to produce results. Enforcement officials can say that they are going to arrest more drunk drivers, but if they do not do it, the public will soon know that they do not mean it. DUI extra enforcement grants coupled with effective public information and education at the local level should produce a meaningful reduction in alcohol-related accidents that can be further evaluated in the future.

ACKNOWLEDGMENTS

Special thanks to Lorraine M. Novak, John Kylor, and Harry E. Balmer, all of the Bureau of Safety Programming and Analysis at PennDOT, for their assistance in compiling the data for this paper; Sharon Mehlbaum and Deborah Snyder, also from the Bureau of Safety Programming and Analysis, for typing this paper; and to NHTSA for their support in funding the DUI extra enforcement grants.

Data Needs for the Operation and Evaluation of New York State's Special Traffic Options Program for Driving While Intoxicated (STOP-DWI)

CLARENCE W. MOSHER

ABSTRACT

The traffic records system developed by New York State in response to the Federal Highway Safety Act of 1966 met the basic needs of the 1970s. However, it does not provide the detailed data needed in the 1980s for evaluation of major safety programs. By using the original traffic records systems as a base, New York State is developing a complex, multilevel, multiagency records system to collect data for evaluation of its Special Traffic Options Program for Driving While Intoxicated (STOP-DWI). This system makes maximum use of data from existing systems administered by state, county, and local agencies.

NHTSA, from the inception of the 402 Highway Safety Programs in the 1960s, recognized the necessity of a uniform traffic records program that was reliable and verifiable in each of the states. The system would need to be established and fully integrated to assess the relative impact of the various countermeasures undertaken in each of the other program areas in each state. As a result, the system was heavily reliant on crash-generated information and would facilitate before-and-after intervention studies that would measure the success of each program.

The thrust of the program as such was adequate for programs in the 1960s and 1970s. However, the broad-based information network necessary to provide both baseline and intervention measures for the major programs of the 1980s is not adequately covered by the traffic records systems established one or two decades ago. NHTSA has highlighted program evaluation for alcohol countermeasures and for restraint use as priority programs for the current administration. The technology necessary for such program assessment involves data retrieval systems of a complex nature, with accident records as just one portion of the whole.

The federal government realized the limitations of its own request of the states to uniformly update the traffic records capabilities when it established its Highway Safety Program guidelines. The standards it established were only the basis for what would naturally follow in succeeding years.

Four classes of information, most of which may be obtained routinely at state or local levels, comprise the data base for all aspects of a coordinated federal, state, and local traffic safety program. This information falls into the following sections: (a) data pertaining to drivers--their licensing, violation records, and financial responsibility; (b) vehicle data, such as make, model, and serial number; (c) highway data on a milepost basis on bridges, structures, tangents, curves, intersections, and traffic control devices; and (d) collision data linked to the drivers involved in accidents, vehicles, and highway locations (1,p.1).

The overall purpose of a traffic records program is perhaps best summarized by the House of Representatives Report (N. 1700 89th Congress, 2nd Session, pp. 10-11):

> Uniform, complete, and accurate accident reports, stored in one center in every State, subject to rapid retrieval and analysis and compatible with a national record system at the Federal level, can tell us not only how many accidents we have, but what kind of accidents they are, where and when they occur, their physical circumstances and the people, injuries, death and damage they involve, what emergency services and enforcement agencies responded and how, and what judicial actions resulted, to mention only the most obvious possibilities.

The role of the traffic records program itself, as the keystone for the entire highway safety effort, was described in that report as follows: "There is no other part of the State program as basic to the ultimate success, nor as demanding of complete cooperation at every jurisdictional level." The report goes on to state: "The effectiveness of the Traffic Records Program is its ability to produce the information needed to support decisions for effective management of the total traffic safety program."

The system itself was designed primarily to assess the magnitude and volume of the highway traffic accident problem on a state and local scale. As such, the traffic records system would identify short-term changes and long-term trends in the magnitude and nature of traffic accidents. It was believed that the traffic records system would provide salient information on high-accident locations and establish causal relationships in accident data. Further, it would assist in the assessment of behavioral factors contributory to an accident, and as such lay the groundwork for the development of countermeasures and for evaluation of effectiveness.

The federal government's guidelines included cautions that are applicable to these data. That is, that the information gathered must be compatible and, at the same time, not duplicative, regardless of its source both at the state and local level. In addition, the concern was expressed (in Report N. 1700) that "adequate and accurate information for reliable statistical analysis must be available to assist State and local officials in safety program planning, prioritization, implementation and program evaluation." It was and is important that the traffic records community remain aware that the agency that contributes information to the traffic records system may in fact be the user of other information from that system at a later date.

The cautions that NHTSA first voiced in the late 1960s have become the watchwords for program development in the 1980s. Specifically they stated (1, Section IV, p.2): "In addition to the data inputs from a multiplicity of operating State and local agencies, each with its own functional objective, mode of operation, and jurisdiction, the Statewide traffic records system must provide for bringing all of the diverse inputs into mutual compatibility. It also must provide for the necessary outputs required by the user groups."

The federal government expressed concern that the information be reliable. That is, in an accident situation, regardless of the reporting system used, researchers have to be assured that data were being gathered for the same drivers, in the same vehicles, reported by the same police, at the same location, at the same time.

It was also recognized by the federal government that many subsections of the overall system could be administered by and the responsibility of several different agencies. The entire scope of the system outlined by the federal government is shown in Figure 1.

Currently, because of a series of budgetary constraints and program evolutions, the number of nationally recognized traffic record program initiatives has shrunken. The importance of traffic records as a cornerstone of Highway Safety Programs is clearly recognized. However, the information needed to plan, analyze, and evaluate highway safety countermeasures far surpasses information contained in the traffic records files of most, if not all, states.

The national priorities of alcohol (drinking and driving) and occupant restraints involve issues beyond simple crash-reported information. Analysis

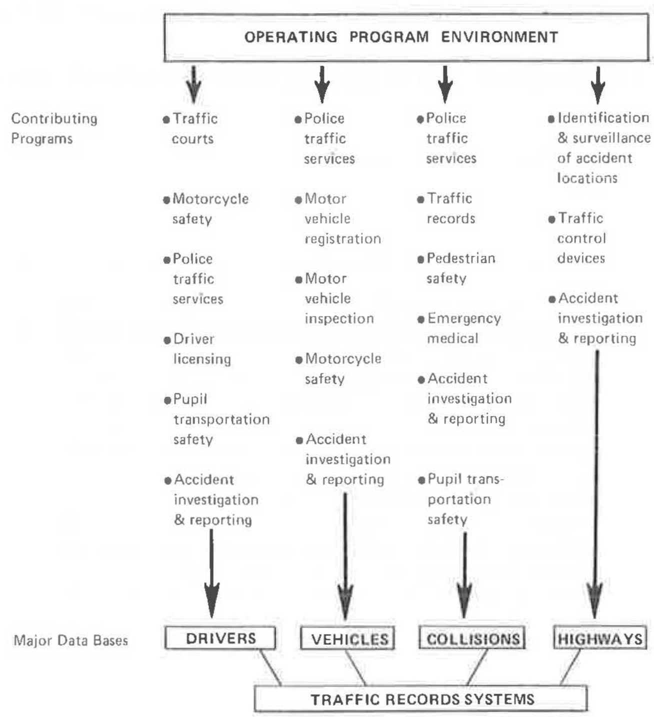

FIGURE 1 Operating program environment.

must be of a scope and nature that is global to the entire traffic safety community and its environment. In any such analysis many questions and issues must be addressed. For example, is the individual who is not properly restrained and becomes involved in an accident representative of the overall driving population? How does the traffic records system account for the drivers who are drinking who do not get involved in accidents? Questions such as these must be examined in any broad-based analysis of highway safety issues.

TRAFFIC RECORDS SYSTEMS AND NEW YORK STATE'S STOP-DWI

In New York State the implementation of the Special Traffic Options Program for Driving While Intoxicated (STOP-DWI), which implemented some of the most significant recommendations of the Governor's Task Force on Alcohol and Highway Safety and a Special Senate Task Force on the same subject, caused the traffic records and highway safety community of New York State to address the issue of adequate and accessible highway safety data.

The STOP-DWI law, which became effective in November 1981, gave county governments in New York State the ability to develop programs especially designed to affect the drinking driver at the county level. Under this law the county's programs would be user funded. That is, the convicted drinking drivers' fines pay for the program. The state has the responsibility for monitoring local program operation, providing technical assistance to localities, and evaluating the administrative elements and impact outcomes of the program. Early in the development of the STOP-DWI program it became evident that New York's traffic records system, prepared in compliance with Standard 10 and the NHTSA Design Manual for State's Traffic Records System, was not comprehensive enough to meet the alcohol and highway safety needs of New York State. That system set the stage but could not meet all the requirements for full program evaluation of STOP-DWI. (A copy of Article 43-A of the New York State Vehicle and Traffic Law, which sets forth the broad scope of the program, is available from the author.)

The empowering STOP-DWI legislation mandated that a full evaluation of the program be made before March 31, 1985. This evaluation effort has been developed to address both the administrative elements of the program as well as the impact on accident and injury statistics in New York State.

Administrative elements of the program include such factors as number of arrests, length of time between arrest and disposition, efficiency of processing individuals through the system, or change in volumes in probation or treatment case loads. Generally, administrative factors define measures of activity and efficiency in each of 58 programs comprising 62 counties in New York State (57 counties plus New York City, which represent 5 counties).

The evaluation of administrative and impact elements of the STOP-DWI program requires an analysis of baseline data for calendar year 1981 (a full year before inception of the law) as well as from November 1981 onward. Many county STOP-DWI program coordinators indicated that they would have little difficulty providing the operational data on a quarterly basis, acquired after initiation of the new law in November 1981. However, many of the coordinators indicated they would have great difficulty in gathering data from periods before initiation of their program. Information was required on 1981 arrests, adjudication, treatment, education programs, and public information programs. (The baseline reporting forms used for 1981 are available from the author.) That kind of information, except for arrest and disposition data in the 10-county Traffic Safety Law Enforcement and Disposition (TSLE&D) system area, was not available in any state-level centralized file.

In addition, although the accident files in the New York State Department of Motor Vehicles (DMV) provided the main data source for impact evaluation, changes in numbers, patterns, and types of accidents and other variables had to be examined to determine their relative relation to the changes in accident statistics. Variables examined included changes in the economy and the unemployment rate.

The New York State Department of Labor provided statistics that, when correlated with the accident picture, indicated a general negative correlation. This tended to support the papers presented at the New York State Association of Traffic Safety Boards' (NYATSB) Conference in September 1981, which first established this correlation. However, relating the full impact of program performance to the DMV accident file and to related data files from other agencies only represented a first step.

The previously mentioned TSLE&D was originally established in a 10-county area in New York State (Figure 2). This system was established to meet the requirements in the Traffic Law Enforcement and Adjudication (TLE&A) component of the state design manual. These 10 counties have the unique capability to examine complete arrest and disposition information. A uniform ticket is issued that can be traced from that point until final disposition. All relevant statistical information on the time frames from arrest to disposition, and final charge at disposition, can be categorized by county, police agency, and court, so that patterns of adjudicative practice can be studied. This information may be integrated with the accident records program for those counties to establish what correlations, if any, may exist between patrol or arrest and accidents and adjudication in terms of sentencing as it relates to arrest.

Although 13.2 percent of the arrests and 12.3 percent of the convictions in the state occur within the TSLE&D area, these 10 counties account for only 7.4 percent of the state's population, 8.6 percent of the state's licensed drivers, and just more than 14 percent of its roadway miles. This records system is extremely useful, and represents an important resource to answer a broad spectrum of alcohol and highway safety questions. However, even a system as comprehensive as TSLE&D provides only a portion of the information needed for the statewide evaluation of STOP-DWI.

The evaluation of the STOP-DWI program is multileveled and multifaceted. It is expected that internal measures of consistency will be developed to assure validity of information through the establishment of mathematical models. The information needed for program evaluation relies on a great deal of data that will be obtained from many state agencies and 58 local sites. The precautionary notes contained in the Highway Safety Program manuals written 15 to 20 years ago still apply today. That is, the data must be consistent, and each contributor must also be viewed as a potential user of the system. However, because of the greater levels of complexity of data, because of greater amounts of data in state and local files, and because of inconsistencies of format between many files, the guidelines established at that time require significantly more resources to provide for consistent and correct analyses of activities and trends.

The New York State traffic records system was adequate to meet the needs of the early 1970s. It does not fully address current needs for data analy-

FIGURE 2 Ten-county TSLE&D area.

sis required by the STOP-DWI law. The current data needs relate to virtually every segment of the local alcohol highway safety system in each county in the state. Because of the legislative mandate under which the DVM is operating, the Department is attempting to establish a broader data system that will indicate, in discrete and measurable terms, STOP-DWI activities related to enforcement, adjudication, prosecution, education, public information, rehabilitation and treatment, and program administration. Carrying out this analysis will require acquisition of data from many local- and state-level sources. Clearly, no single existent traffic records system contains all elements required for such a broad analysis.

STOP-DWI EVALUATION FRAMEWORK

Now that the STOP-DWI program has moved into its second full year of operation, there are attempts to identify and address deficiencies in the state and local information network. As a first step, the Office of Alcohol and Highway Safety (OAHS) identified six areas that generally reflect program activity throughout the alcohol and highway safety system: demographics, accident data and blood alcohol content (BAC) data, arrest (enforcement) data, conviction and disposition (adjudication and treatment) data, education data, and public information data.

OAHS staff proposed that these six data categories would provide the basic framework against which (a) specific program activities could be compared and analyzed, and (b) overall program trends could be identified over time. The scope of each area, as well as perceived deficiencies, are noted in the following sections.

Demographics

Demographics data include (a) population data from New York State through the federal Bureau of the Census, (b) population data from the New York State Department of Commerce, (c) road miles traveled as measured in miles of centerline roadway by the New York State Department of Transportation, (d) number of licensed drivers by age and sex and by county from the New York State DMV, and (e) number of registered vehicles by county, also from the DMV.

In addition, integrated into this model are the related local highway safety grants, either ongoing or just completed, which will potentially affect program results. This information is submitted to the OAHS by the Governor's Traffic Safety Committee (GTSC). Such information is useful in assessing activity in specific counties. For example, if a county is receiving sizable Section 402 funds for an enforcement program that will at some point overlap the STOP-DWI enforcement effort, the impact of the Section 402 effort must be accounted for.

Finally, at various stages in the program surveys are being taken to assess public opinion, knowledge, and perceptions. Ideally, public surveys should have been administered before initiation of the STOP-DWI program. Because this was not done, surveys conducted in New York State have attempted and will attempt in future scheduled surveys to ascertain perceived changes in knowledge and attitudes about drinking and driving.

The total body of information from these several sources will help to define a general demographic profile of each county.

Accident Data and BAC Data

Accident data and BAC data provide for a specific measure of alcohol and highway safety activity at a county and statewide level. Accident data and analyses obtained through standardized accident reports are now being integrated with coroner's reports on BAC and Department of Health reports on morbidity and mortality. Other than the Department of Health's contribution, the accident analysis of trends is accomplished as it is recommended in the traffic records program manual and integrates all recommended portions of the system. However, New York State's data, like other states, has consistently indicated an underreporting of alcohol involvement. When the STOP-DWI program began, OAHS observed an increase in enforcement training efforts and public awareness of the issue. It has been suggested that these two factors helped bring about more accurate reporting of alcohol-related accidents and of actual BAC levels.

Arrest Data

Arrest data provide a significant indicator of alcohol and highway safety enforcement activity. The computer files at DMV contain a relatively complete conviction picture in the state. However, this file may not accurately depict actual levels of arrest activity because of the possibility of reductions or dismissals. Since implementation of STOP-DWI, many questions relating to arrest activity in New York State have been raised. As a primary measure of the program's activities throughout the state, there was a need to ascertain if alcohol-related arrests were increasing, and if so, at what rate. Such factors as time of arrest or police agency were also of interest. In addition, more sophisticated questions regarding arrest activity and potential for accident involvement have been raised. Arrest data to answer such questions proved to be only fractionally available.

In New York State the State Police account for approximately 15 percent of the arrests. They do not patrol in New York City and they have their own record keeping system. The Division of Criminal Justice Services' (DCJS) Bureau for Municipal Police is responsible for the aggregation and analysis of arrest information from each police agency. However, DCJS must wait for reports to be filed by the local police agencies in the state. DCJS does not at this time have the personnel to fully verify the accuracy of data currently being reported to the Federal Bureau of Investigation (FBI). In addition, their file is primarily based on fingerprintable offenses that would include driving while intoxicated (DWI) (BAC of 0.10 and above), but not driving while ability impaired (DWAI) (BAC between 0.05 to 0.09) cases. To further complicate the issue, the DCJS system is not tied into the DMV accident file. As a result, there is no assurance that multiple reports relate to the same event.

Each of the 58 STOP-DWI coordinators have been able to report on arrests in their counties as they received them from local police jurisdictions, if indeed they received them. It is believed that this direct reporting procedure will provide a somewhat more accurate arrest picture in each county than is currently available until TSLE&D is implemented statewide. Often there is agreement between the arrest files maintained by DCJS and the county-submitted arrest data. But there can be a discrepancy of as much as 25 percent that, in larger counties, may represent approximately 2,000 cases.

Although the importance of arrest data is recognized, the availability of such data in a timely and accurate format is difficult to ascertain on a statewide basis. The 10-county TSLE&D area is the only area in New York State that can provide a complete and accurate arrest picture. For the rest of the state, OAHS must rely on statewide files and individual county reports to approximate arrest activity.

Conviction Data

Conviction data provide for a summary of disposition of DWI and DWAI cases. Files maintained at DMV contain all such data reported by all the courts in the state. These files exist for the purpose of imposing legislated and regulated penalties for alcohol-related convictions. Data contained in the files describe fine levels, jail sentences, and recidivist activity. Although it is believed that this file is reasonably accurate, there are some deficiencies in its data as well.

Except for the TSLE&D system in the 10 demonstration counties and the administrative adjudication system operated by the DMV in New York City, Buffalo, Rochester, and Syracuse, the state is totally reliant on the court of conviction filling out a form and submitting it to the DMV. It is admitted that some judges have not reported in years. The DMV does record actions against a driver's license, but any court activity that does not result in a conviction is lost. OAHS is able to secure that information only from TSLE&D files and from the State Police, which follow all tickets from issuance through disposition. All other activity is for the most part lost unless the county STOP-DWI coordinators can provide reports on court activity in each county.

The specific interventions of probation or rehabilitation are important components of the state's conviction data. Convicted drinking drivers represent the largest single population of individuals referred for probation. Likewise, a significant number of clients mandated for alcoholism treatment come from the DWI-convicted population. Information on individuals placed on probation or placed in alcoholism treatment as a result of DWI represents an important data element within New York State's total system. Part of this information exists in the DMV conviction file, but the majority of such records are housed in files maintained by the Division of Alcoholism and Alcohol Abuse (DAAA) and by the Division of Probation. Identifying the history of the individual who is arrested, convicted, referred to probation, goes to a drinking driver program (DDP), and is referred to treatment requires access to files in at least five different agencies. The job of assuring that OAHS is following one individual throughout that system is a difficult, if not nearly impossible, task.

The basic data in the conviction file provide a general review of fines and penalties imposed after conviction. Specific information relating to such interventions as probation or alcoholism treatment requires significant cross analysis and validation between multiagency data systems.

Education and Public Information Data

The other two components defined for the evaluation are education and public information. Although it is possible to ascertain how many children are now receiving alcohol and highway safety education, tracking these individuals through a lifetime of driving is again difficult, if not impossible. In addition, assessing the relative merits of one method of education versus another has been debated for years. It is possible to count the number of public service announcements, the number of public speaking assignments, and the number of articles on accidents and arrests, but to assess how this affects driving is problematic.

Although the New York State Department of Education maintains substantial amounts of data on such items as school enrollment, fiscal reimbursement formulas, and general levels of academic achievement, specific measures of alcohol and highway safety education are again difficult to obtain, at least as OAHS begins to look for and plan to incorporate this type of data in its evaluation system(s).

Assessing the overall impact of public information activities presents similar problems. For example, although the number of public speaking activities or the number of newspaper articles devoted to alcohol and highway safety in a locality can be itemized, there are still significant problems in assessing the impact of any or all of these activities.

SUMMARY: CURRENT SYSTEM AND LIMITATIONS

Implementation of the New York State STOP-DWI program provided for the implementation of a statewide, multidisciplinary set of program interventions. Such a level of highway safety program activity was virtually unprecedented in such a short amount of time. The mandate to the DMV to carry out a thorough and comprehensive evaluation of the program has highlighted the need for a responsive and accurate highway safety data system.

The traffic records system put into place in New York State in the early 1970s provides a basis for broad analysis and general study. However, many data elements, other than those noted categories, do not exist in any one location. Files that define county demographics come from nonuniform data systems. Files that reflect accident and BAC information exist but rely on accurate coordination with Department of Health files on morbidity and mortality. In addition, accuracy of such files depends directly on accuracy of source documents filed by enforcement offices. Arrest data are only as accurate as source documents submitted by appropriate police agencies. Except for data in the 10-county TSLE&D area in New York State, arrest files are compiled by submissions to the Bureau of Municipal Police in the DCJS. If these data are missing or inaccurate, there are few options at the state level to establish a complete file. Accuracy of record keeping at the local level, apart from the TSLE&D area, is not guaranteed.

Conviction data exist in the centralized DMV's file. However, evidence of dismissals or reductions is unavailable from that file, and only available from TSLE&D, as is information on such specific interventions as referral to probation or rehabilitation.

Consistent information on activities related to education and public information programs is among the most difficult to obtain. Centralized data provide for only the most cursory review of county-level education incentives. The impact of public information efforts is likewise difficult or impossible to ascertain from any current state or local data system, and must be developed.

FUTURE DIRECTIONS

Although the preceding summary of data needs and data availability may appear somewhat discouraging, DMV has implemented several initiatives that, it is believed, will address the evaluation needs of the STOP-DWI program.

During the past year analysts in the OAHS have established working relationships with many other agencies to begin to acquire and analyze data from other files. Although such data are often in formats different from DMV's file formats, the Department has begun the process of verifying and enlarging on county-specific data files. County coordinators all across the state have begun to contact their local constituencies to recommend accurate and timely submission of data to appropriate state agencies.

Possibly the most important activity regarding accurate data acquisition has been implemented since passage of the STOP-DWI law. The OAHS in the Department of Motor Vehicles has developed a comprehensive data report that is intended to fill in, to as great a degree as possible, perceived deficits in the alcohol and highway safety information system. The Administrative/Impact Evaluation (AIE) forms have set forth an information reporting system that will provide discrete measures of activity throughout the local system. These forms require a local STOP-DWI coordinator to acquire directly at the local level significant amounts of data on specific elements of the local system. Data have been requested for the baseline year of 1981 and for each quarter of subsequent years. (Copies of the quarterly STOP-DWI reporting forms are available from the author.)

The OAHS believes that the AIE reports, submitted for each county, will provide a complete description of pre- and post-STOP-DWI activity. A cursory review of the forms indicates that a great amount of data is being requested and secured. In some cases the OAHS knows that local data will not be available. In that event OAHS will attempt to provide as much information as possible from state files, while acknowledging their limitations.

Despite the obvious shortcomings, a complex multilevel, multiagency records keeping system is slowly coming into place. Will it answer all of the questions? No, not immediately, and perhaps it never will to OAHS's satisfaction, but it has come much closer to understanding alcohol's effects on highway safety. The same type of expanded information network may be essential to assess the impact of occupant restraints. An accident-based system as promulgated in the 1960s and 1970s, which provides the basic building block and which must be in place, is just that, a building block. It can give a portion of the picture of what happens on the roadways. But without the remainder of the components in place in a verifiable and reliable manner, the degree that an accident is representative of the entire highway safety community is at best a matter of an educated guess.

To answer the questions arising from an informed and knowledgable constituency, an expanded system of highway safety records must be developed and integrated in each state. Of course, each state must assess its own data subsystems and their ability to

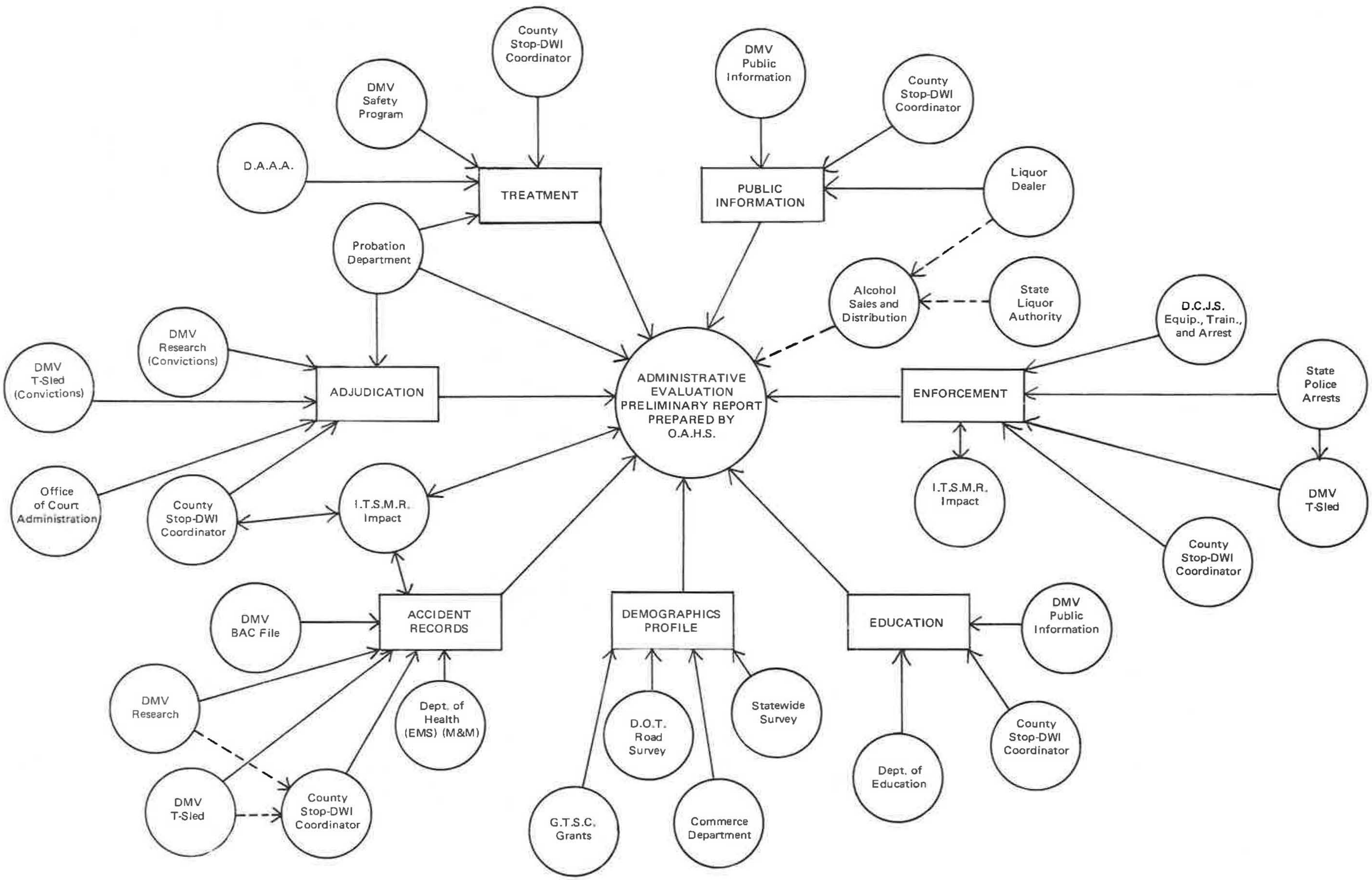

FIGURE 3 Data-acquisition model.

be integrated into a comprehensive system before the implementation of new complex highway safety programs. Without reliable and valid baseline data created by all the agents who are potential users and contributors to the system, accurate measures of success will be, at best, difficult.

In New York State the Office of Alcohol and Highway Safety in the Department of Motor Vehicles is attempting to develop a complete data system in two general ways. First, OAHS is building on the foundation of the original traffic records system put into place years ago by integrating in consistent ways data from other agencies. Second, OAHS is requiring that each county coordinator submit detailed and accurate reports on all appropriate county-level alcohol and highway safety activity. In this way OAHS is using the best data available, either from the state or local systems, to carry out a comprehensive evaluation of the STOP-DWI program.

OAHS believes that the total data acquisition model (Figure 3) will provide the most complete and accurate picture of alcohol highway safety activity in New York State. Although this is not yet an ideal system, it is believed that the evaluation model and the data-acquisition procedures put into place will provide the best possible basis for the program assessment that must be provided to the Governor and the Legislature on March 31, 1985.

REFERENCE

1. Highway Safety Program Manual, Traffic Records. Volume 10. NHTSA, U.S. Department of Transportation, March 1975.

Meeting the Challenge of Traffic Information Systems in the 1980s

MARK H. LARRATT-SMITH

ABSTRACT

The Transportation Regulation Program of the Ontario Ministry of Transportation and Communications (MTC) uses five distinct, but interdependent, information systems: the Driver Licensing and Control System, the Vehicle Registration System, the Motor Carrier Performance and Enforcement System, the Traffic Accident Information System, and the program's internal Management Information System. The Vehicle Registration System has recently been revised in response to pressure from the public, police, and courts. To avoid future massive catch-up projects dictated by client dissatisfaction and as a response to growing external demands and pressures placed on MTC's information systems, several initiatives have been adopted, including (a) establishment of a Systems Improvement Office that oversees system maintenance and improvement, (b) development of priorities for system activities, and (c) career training to familiarize all managers in MTC with the operation of information systems.

The term traffic information systems is frequently defined in narrow terms to refer simply to data files that contain information concerning traffic volumes or accident information. As in so many other areas of government activity, the growth of information technology and the demands on the management of information systems that flow from it have inevitably rendered this narrow view of traffic information systems obsolete.

By using the perspective of an organization with a range of responsibilities that includes all highway users in the province of Ontario, it is proposed that, for purposes of this paper, the term traffic information systems be redefined to include all user-related data (excepting only that which is primarily related to the infrastructure of the highway system) as a prelude to discussing the challenge of the management of such systems in the years ahead.

Within the Transportation Regulation Program of the Ontario Ministry of Transportation and Communications (MTC), five distinct information systems have been identified: the Driver Licensing and Control System, the Vehicle Registration System, the Motor Carrier Performance and Enforcement System, the Traffic Accident Information System, and the program's internal Management Information System.

The Driver Licensing and Control System encompasses the entire process of gathering, storing, and retrieving information about Ontario's 5 million licensed drivers, including the control and suspension components related to convictions, demerit points, medical impairments, and nonpayment of fines.

The Vehicle Registration System includes all aspects of the collection, storage, and retrieval of information concerning the 5.2 million vehicles registered in the province of Ontario, including associated taxation collection and audit components, police interfaces, mechanical fitness requirements, and certification of valid insurance.

Although many of its elements have existed in manual form for many years, the Motor Carrier Performance and Enforcement System has only recently been defined as a coherent and distinct system that is still in the process of being elaborated. In concept, it contains data related to all operators of commercial motor vehicles who use Ontario's highways. In the future this system will include the classification of carrier, information on vehicles operated, and a wide range of reported infractions from safety violations through labor law, weight and dimension, registration, and operating authority convictions.

As in many other jurisdictions, the Traffic Accident Information System has existed in Ontario for many years, but it is suspected that the MTC is by no means unique in its relatively unsophisticated manner of using this system, which is, after all, the fundamental base on which all highway safety programming must rest. In Ontario considerable effort has been directed to the development of usable information related to infrastructure safety (with significant safety results), but little has been done to develop a sophisticated method of using accident data to direct safety programs that are aimed at the driver and the vehicle. This omission is particularly significant in that current research indicates that more than 80 percent of all accidents are attributable to driver error.

The program's internal Management Information System is also in a state of considerable change. The pressures that will be elaborated on in this paper have resulted in the expansion of traditional personnel and financial short-range budgeting and control information into a long-range planning and control system encompassing not only input information but also information on products produced, services rendered, and results achieved.

As will be apparent from even this brief description of the five general systems that are included in the broad definition of traffic information systems, there is considerable interdependence among all five. The driver system incorporates accident-related data; the carrier system is being built on commercial vehicle registration data and in the future it will increasingly need to incorporate relevant driver information; and the Management Information System monitors all of the resource inputs and product outputs relating to the other four systems.

With this broad definition of traffic information systems, it is evident that the management challenge is a formidable one. This challenge is examined by first looking at the external environment that affects the traffic information systems. Next a specific case, which involves the recent redesign of the entire Vehicle Registration System in Ontario, is discussed as an example of the issues raised and the lessons learned from attempting to manage change in a large system under many of these external pressures. Finally, a review of some specific initiatives that Ontario has under way for the future as a result of the lessons learned is presented. This information is presented, so far as possible, in a generalized fashion in order that it may be of maximum assistance to other jurisdictions facing similar challenges.

EXTERNAL ENVIRONMENT

There are many external demands and pressures placed on government transportation departments to maintain accurate and effective information systems. These demands and pressures increase with the growing sophistication of data systems. One example of the growth in sophistication of data systems in Ontario is evident in the recent decision by MTC to manage the denial of license renewals for nonpayment of motor-vehicle-related fines. Although MTC had previously acted as an enforcement agency for the courts on driver licensing issues, it has only recently begun to do the same for vehicle licensing.

Another example of the growing complexity of data systems in Ontario is the increasing diversity of accident statistics related to types of vehicles. With the advent of new definitions in heavy truck and other vehicle categories, maintenance and analysis of accident statistics become more complex, and even more important for policy-related purposes.

In addition to the growing complexity of the content of systems, departments of transportation must also contend with the changing technologies of data systems. The growth of data records and systems encompasses new information technologies such as microprocessors and distributed processing, all of which are in a constant state of flux. Managers of systems must therefore continually revise systems technologies as well as systems content to keep pace with information flows.

Another environmental challenge to any traffic information system involves financial constraints that create pressures to find more efficient means of maintaining systems. The heavy pressures for greater organizational efficiency tend to feed the automation pressures, because the obvious, if at times superficial, attraction of automated technology is the perception that it will save organizations money and personnel. Too often this is not the case, and in exchange for a saving in clerical staff an organization finds that it has been burdened with a fragile system that is expensive to maintain and whose shortcomings are extremely evident to the public.

Another factor that is closely related to financial constraints is the increasing pressure on public agencies to justify the money that they are given. This pressure for increased accountability for decreasing funds is pushing public agencies heavily in the direction of program effectiveness evaluation, so that their activities can be either justified or eliminated.

The term "program evaluation" finds expression in various ways in different jurisdictions, with "value for money auditing" being one particular term currently in use in Ontario. Value for money auditing provides for the evaluation of all government programs against their ultimate public benefit. Any such system is bound to include a high degree of subjectivity, but the exercise does stimulate a useful emphasis on the development and tracking of appropriate numeric indicators whose linkages to dollars spent, however soft, do tend to focus the debate on issues of public benefit. In this respect the subject of highway safety presents a fascinating challenge. Accident causation is an enormously complex subject, but focusing on linking programs to real-world results in terms of accident reduction has the potential of being a useful device in a field where the choice of programming initiatives has frequently been totally unrelated to any real analysis of potential benefit.

A final external pressure that affects the management of traffic information systems is the issue of fundamental rights. In Canada a Charter of Rights, which is somewhat analogous to the American Bill of Rights, has recently been adopted. In this respect Canada is just now starting to face issues that have been of concern to Americans for nearly 200 years. Although the timing may be a little different in this one respect, it is believed that the pressures associated with the adoption of the Canadian Charter of Rights and also with the extension of human rights legislation in Canada are similar to those that exist in the United States. The right to access information currently maintained by a public agency can create substantial costs and administrative complexities. In addition, freedom of information initiatives raise questions as to whether or under what conditions information should even be collected in the first place. In turn, these questions place additional pressure on justifying government programs that require the data collection, in terms of their impact and effectiveness from a broad public benefit point of view. The debate risks becoming circular when agencies resort to collecting additional information to prove the effectiveness of the programs that required the information that initially came under scrutiny.

Also within the context of rights and privileges, the issue of freedom of information poses an additional constraint on systems management. Newly imposed rules for public access to information in many jurisdictions place government in the position of having to maintain tightened confidentiality on many items, whereas making freely available information that in the past has never been released. This greatly increases the complexity of the management task related to the collection, segregation, retrieval, and production of information.

ONTARIO VEHICLE REGISTRATION SYSTEMS PROJECT

Turning from a general outline of the external factors that pose a challenge for the management of traffic information systems, a recent Ontario experience is presented that illustrates some of the practical difficulties involved when dealing with some of the broad environmental challenges previously discussed.

The Vehicle Registration Systems Project (VRSP) was formally completed on March 31, 1983, although certain details of implementation have carried on since that date. It has involved development of a new plate-to-owner system of license renewal; a staggered renewal of passenger and light commercial vehicles by owner birth date or system generated date; a complete on-line system to more than 300 private agents across Ontario; a revised fee structure, which is fee to plate with a flat fee; a turnaround document that is mailed to vehicle owners in advance of their renewal date; and the capability to deny registration renewal of the offending or replacement plate for nonpayment of parking violations. The project itself cost approximately $11.8 million and involved the complete rebuilding of the previous vehicle registration system. It has been the largest systems project ever undertaken by the Ontario government and the biggest motor-vehicle-related project in North America in recent years.

MTC became involved in VRSP in order to replace a poorly operating automated batch registration system. The inaccuracies and especially the delays in recording transfers on this system caused a high degree of frustration for the police community and the courts. In the case of the police, the system did not meet the requirement that a license plate should provide an accurate pointer to the owner of a vehicle. This pointer is a major investigative tool for all police work, not just for highway-traffic-related law enforcement. At the same time the court system wanted an increased requirement for an accurate identification system that could permit denial of registration for nonpayment of parking fines by vehicle owners as an alternative to the expensive and unpopular process of executing summons and ultimately putting recalcitrant offenders in jail. Although not a primary motivator in undertaking the project, the unhappiness of the general public over long annual lineups for registration renewal was also a factor in the planning of the project.

What is significant about this description of the genesis of VRSP is that, in effect, MTC had lost control of the system. It was the users and their unhappiness with the system that became the generating force for change rather than the organization's own initiative, including its determination of what it could feasibly handle without major disruptions. In retrospect, this loss of control before initiation of the project had extremely important effects on the costs and complexity of the project itself. It has also led to a strong determination by MTC that, through anticipatory planning and attention to user needs, the Ministry will ensure that it never again enters into a massive catch-up project where client dissatisfaction leads to dictated solutions.

Another important and related aspect of this project was that the external pressures involved were related entirely to service improvement. The simplistic notion that automation saves money became a factor in many people's mind, particularly in the central agency responsible for budget control. This not only meant a continual need to explain what indeed the program was about, but it caused difficulties for everyone when severe financial constraints within the Ontario government coincided with the final implementation of the system, which was then fully committed and which involved expensive new services (many of which were impossible to cost-out in advance).

VRSP was implemented over a period of 4 years and involved approximately 40,000 person-days. On the whole it has been a successful program, in that it was implemented on schedule with relatively little negative impact or adverse reaction from the public. It is, however, a qualified success because the system is still not fully in place in terms of some behind-the-scenes adjustments that are still being made. In addition, the cost estimates have grown significantly from the beginning of the project, and they are well above the expectations of the agencies involved.

A number of the lessons learned from VRSP are useful in a discussion of the challenges facing managers of traffic records in the 1980s. The first lesson that emerged from this particular project was the notion that an information system of this type must clearly identify its users and their needs. It must do this on a regular and ongoing basis. It must also set priorities for those needs and clearly articulate what can and cannot be achieved within identified time and dollar constraints. Finally, the information system must force clients, particularly other government agencies, to participate fully in a justification of the public benefits of their wish lists.

Another rather important lesson that MTC has realized from the VRSP experience is the need to build total systems expertise into any organization that is managing traffic records. As automation becomes an important part of information systems, organizations appear to progress through a number of identifiable evolutionary stages. Typically, this evolution starts with a fascination with the hardware itself. Once the computers have been installed, preoccupation gradually shifts to the programming task. At this stage the software experts become in effect the "high priests" of the system. A third stage has become apparent recently with the trend toward so-called user friendliness, which has diminished the need for highly specialized systems programmers within organizations. This stage has expanded the focus to broad-based systems analysts who can deal with issues relating to the overall management of systems and particularly with the people components that are, after all, the most complex and difficult to manage. The final stage of this evolution is to regard the entire operational management team as part of an organization's systems expertise, so that there is not a gulf placed between those who run the system and those who alter it.

At the outset of VRSP, MTC had not progressed much beyond the second or programmer preoccupation stage. It was discovered that the organization had a dirth of broad systems expertise and that it tended to underestimate both the capability of the operational people to do systems work and their indispensability in keeping the previous system running while the new one was being built. As a consequence, the Ministry became more dependent on highly specialized consultants than would have been desirable. Although consultants are essential to any large-scale systems project because of their ability to provide a pool of highly specialized talent on an ongoing basis, they need to be contained within a strong and tightly controlled organizational matrix to ensure that their talents are appropriately directed and that the results of their work are appropriately transferred to operational staff.

Another important lesson arising from VRSP involves the need for a much clearer understanding of the critical role that the selection and management of a systems project methodology plays in the success of a project. For VRSP, the Ontario government used an existing technology entitled Spectrum. Although the methodology provided many benefits to the project, in the end it proved inadequate. It tended to focus unduly on the production of detailed information rather than placing emphasis on critical decision-making points.

In addition, Spectrum proved to be linear in concept, whereas recent developments in systems technology, as well as the pressure of meeting imposed deadlines, led MTC to adopt an iterative or interactive approach to project tasks. In effect, because of the delay of certain key policy decisions, much of the detailed design, programming, and testing took place virtually simultaneously. One of the specific conclusions that emerged from this project, therefore, was that for any large systems project, time should be taken to devise a custom methodology, or at the very least, to carefully and thoroughly adapt an existing methodology to the requirements of the project.

ONTARIO'S RESPONSES TO ENVIRONMENT AND EXPERIENCE

Given the Ministry's experience with VRSP and its assessment of the environmental challenges facing it in the latter 1980s, MTC has devised a number of initiatives to deal with the future management of information systems.

One such initiative has been the formation of a Systems Improvement Office. This group, which is distinct from the government service organizations that supply computer time and programming support, reports to the senior operational management responsible for operating all of the systems that have been described in this paper. Its mandate is the maintenance and improvement of these systems under the direction of operational managers. The office currently has a complement strength of 31 positions.

There are two aspects of the creation of this group that should be emphasized: funding and control. With respect to funding, there is an interesting parallel to the issue of preserving highway infrastructure. As in many other jurisdictions, highway engineers at MTC have developed a relatively sophisticated approach to highway maintenance based on the fundamental concept that, to build a road, sufficient funds must be allocated, not just to prevent its physical deterioration but also to make operational improvements that will maximize longevity before there is a requirement for massive and costly reconstruction. This same philosophy applies to traffic information systems. Especially with the advent of real-time on-line systems, such as the new VRSP, it has become imperative to put aside, on a continuing basis, a significant portion of the cost of the development of such a system for its maintenance and improvement. Within the $65 million budget of the Transportation Regulation Program, MTC is allocating $3.6 million (as a minimum base level of funding) to the Systems Improvement Office to provide all of the services related to systems maintenance and improvement that are required by operational managers.

The second point of emphasis with regard to this new office is the issue of control. The first principle here is that the Systems Improvement Office operates on a zero-based budget. In fact, the $3.6 million budgeted for it is allocated among the operational managers responsible for each of the five systems that were identified at the beginning of this paper. To allocate this funding on a realistic basis, MTC has developed a relatively sophisticated method of setting priorities for the needs of the individual systems. For a number of years now, MTC has had a functioning cyclical strategic planning process. One major component of this process is the development of an annual long-range plan with a 5-year horizon. In order to set priorities for systems, MTC has added an additional component to the long-range plan for each of the five systems. These systems long-range plans are developed after the general direction of the program has been well-enough established that they can identify the implications of trends and changes of direction. The systems plans are developed by the Systems Improvement Office under the supervision of a committee composed of the key operational managers and internal users of each system. Once these plans have been developed, they are integrated into a series of priorities by the overall program and the available funds are allocated to each user area to purchase systems support. The user committees are also responsible for generally directing the work of the Systems Improvement Office in undertaking developmental activities that flow from this plan.

A brief summary of current program priorities and attendant systems activities will help illustrate the way in which this approach is used. For the driver system, emphasis has been placed on increasing the sophistication of the driver improvement programs. This means that there will be a high priority placed on improving MTC's capability of tracking problem drivers in order to develop selective treatment strategies that will affect motivation and performance. Improved tracking will also generate effectiveness information for use in justifying either the retention or the alteration of programs that have a high degree of public sensitivity.

Also for the driver system, the introduction of a photo driver's license in Ontario is under serious consideration. This innovation will obviously have significant systems implications. In preparation for such an initiative, MTC has been concentrating its efforts on necessary preparatory improvements to the existing system and on extensive evaluation of user requirements and system alternatives.

The priorities that currently exist for vehicle systems center around the final implementation and maintenance of the new system that has just been put in place. As mentioned previously, MTC is also extremely conscious of the need to stay ahead of demand in this field, so that attention is being given to identifying other areas of potential client demand.

For the Motor Carrier Performance and Enforcement

System there appears to be an enormous potential demand for new information to be collected, maintained, and retrieved in innovative ways. The development of a Commercial Vehicle Operators Registration System, which would identify the operators of all trucks on Ontario highways, irrespective of ownership, lease arrangements, and so forth, has been recommended by groups studying both truck safety and economic regulatory reform. Systems activity in this field is focused at the conceptual stage in anticipation of a major effort.

Extensive plans for the accident information system are currently being developed by MTC. As previously mentioned, MTC has in the past collected a great deal of information concerning accidents without using it to its fullest. The priority with regard to using accident statistics no longer lies with the traffic engineering function, but instead is used for the driver system where improved accident analysis can, it is hoped, identify and justify opportunities for new programming.

Another high priority with regard to the accident information system is in the development of a service capability for the evaluation of safety programming. With funds as scarce as they are currently, both in government and in the private sector, there is obviously a need to provide all those interested in highway safety with better support in assessing whether existing or proposed safety initiatives are effective. As the custodians of the accident system, MTC sees an obligation to assist everyone--from local community groups to industry to enforcement agencies--with data and interpretative help in focusing their efforts. Obviously, MTC's ability to help will be constrained by its own resource limitations, but it is anticipated that even a small start may show tangible results.

The current direction of activity in the Management Information System also relates to concerns about effectiveness. It is commonplace to say that the difference between business and government is that only the former has a bottom line. One consequence of this has been that, over the years, it has been both more difficult and less critical to define the products of government than to identify those of a corporation whose survival depends on their profitability. For years governments have tended to manage their affairs on the basis of inputs rather than on the basis of outputs. One of the beneficial results of the recent constraints on government funding at all levels has been to force changes to this particular form of management. These changes go under a variety of names in different jurisdictions, but the net result has been to focus on the planning and control of government programs against specified outputs as well as against resource inputs. Within MTC there is a project under way that anticipates fully automating the operational planning and control of all the resources and outputs of the Transportation Regulation Program. Although this particular internal management system may not be strictly speaking a traffic information system, it is essential to the management of the traffic information systems previously mentioned.

Thus far the discussion of MTC's response to both the environment and its experience has focused on two areas: the creation of a Systems Improvement Office and the development of a capability for setting priorities for systems activities in the five identified areas.

A compliment to these initiatives is the undertaking of a long-term commitment to the development of systems literacy for all operational managers. The first steps in meeting this commitment involve undertaking the development of a training plan to provide greater familiarization to existing managers and a commitment within the context of a comprehensive human resources plan to ensure that managers who progress through the Ministry's ranks receive direct exposure to systems project work as part of their career assignments.

CONCLUSIONS

In conclusion, there are some key messages that Ontario can relay to other jurisdictions in the business of traffic information systems management. The first of these messages is that organizations should broaden their horizons when examining traffic records systems. These systems are so large and so pervasive that they can no longer be managed simply as specialized data bases that focus on specific functions. To attain maximum benefit from systems, uses should be articulated and should have set priorities as part of a total government approach to problem solving. In addition, operational control and systems literacy are key components of systems management and must receive attention not just in short-term organizational solutions but in an integrated longer-term approach to human resource development. Finally, it is crucial that systems are developed with the ultimate user in mind, so that the community at large may obtain maximum benefit from systems that are developed and maintained at public expense.

On the whole, MTC has had some significant success in managing and upgrading its traffic information systems. Through consultation with other jurisdictions and the development of new systems ventures, MTC hopes to meet the challenges that lie ahead within this volatile field.

Severity of Large-Truck and Combination-Vehicle Accidents in Over-the-Road Service: A Discrete Multivariate Analysis

THIPATAI CHIRACHAVALA, DONALD E. CLEVELAND, and LIDIA P. KOSTYNIUK

ABSTRACT

The severity of large-truck and combination-vehicle accidents was investigated by using 1980 Bureau of Motor Carriers Safety data. The analysis was based on 19,263 accident involvements of such vehicles engaged in over-the-road operation. A two-stage discrete multivariate analysis procedure was used. The first stage involved selecting significant independent variables from 18 candidate variables. The second stage involved modeling accident severity by using two independently estimated logit models, one for the probability of a reported accident being fatal and the other for the probability of a less-severe accident having nonfatal injuries. Differences in the effect of the variables for the four predominant truck types (straight trucks, singles with van, singles with flatbed or tanker, and doubles) led to their separate analyses. The accident severity for each truck type was relatively simple, and most shared common variables and interactions: the main effects of type of collision (single-vehicle or collision with car or commercial vehicle), road class (number of lanes, median separation, and rural or urban setting), and environment (day or night, and road surface wetness condition). The interactions involving road class and environment, and road class and collision type, were usually important. No driver characteristics were found to be significant. Particularly severe accidents were collisions involving passenger cars and doubles, straight trucks, or loaded flatbed or tanker singles on undivided rural roads; collisions involving cars and van singles on undivided rural roads at night; and collisions involving cars and doubles on divided rural roads.

Accidents involving large trucks and combination vehicles are of great concern to the public and to government agencies. The safety issue is more important than ever before because of the 1982 federal legislation that lifted restrictions that long barred the operation of some large combination vehicles in many states. A portion of the findings from a research effort concerning large truck accidents (1) is reported in this paper. The various characteristics that affect the severity of accidents involving at least one large truck or combination vehicle are discussed. A discrete multivariate analysis of a national truck accident data set, including both variable selection and modeling, was carried out.

Why one accident results in a fatality while another does not is the result of a complex interaction that involves the vehicles, the drivers and other occupants, the road, the environment, and chance. Because of such complexity, direct theoretical work has not yet led to sufficient understanding of the problem, and data developed from actual accident experience are analyzed in search of the important determinants of accident severity.

A number of past studies (2-6) compared the accident severity of truck-tractors pulling one trailer (singles) and truck-tractors pulling two trailers (doubles) on the basis of their respective numbers of fatal or injury accidents per million vehicle miles of travel. Hedlund (7), in a study of the effects of road class and vehicle weight on the accident severity of collisions involving combination vehicles and cars, reported that the odds of a fatality varied significantly with the class of roads on which the accidents occurred. Two-lane rural roads were reported to have the highest odds of fatality, followed by four-lane rural roads and urban streets. Hedlund found no independent effect of vehicle weight. No reported studies that simultaneously examine a larger set of possible contributing factors have been found.

DATA

The 1980 Bureau of Motor Carriers Safety (BMCS) data file was used in this research. It contains the records of accidents that involved interstate carriers, subject to the U.S. Department of Transportation Act of 1966 (49 U.S.C. 1655), that resulted in death, injury, or at least $2,000 of property damage. Each accident involvement record includes information on the accident severity, the vehicle (cargo, number of trailers, body style, power unit), driver (age, years with company, hours on duty), environment (light condition, weather), road class (number of lanes, median separation), road surface condition, and accident characteristics. In all, 74 parameters are given for each involvement.

There were 32,245 accident involvements of large trucks and combination vehicles reported to the BMCS in 1980. About 7 percent of these involvements resulted in at least one fatality, 54 percent in at least one nonfatal injury, and the remaining 39 percent in no injury but with property damage exceeding $2,000. Of these, 6,258 accidents involved vehicles engaged in local pickup and delivery service, and the other 25,670 accidents involved vehicles engaged in over-the-road operation. When these accident involvements were broken down by vehicle type, 3,099 involved single-unit trucks (straight trucks), 26,409 involved singles, 1,218 involved doubles, and 1,519 involved tractors only or an unknown vehicle type.

METHODOLOGY

Variable Selection Procedure

The objective of the variable selection was to select a set of significant independent variables from a large number of candidate variables. The variable selection criteria were based on three test statistics--the Pearson chi-square (X_p^2), the total chi-square (Q_T), and the generalized Cochran-Mantel-Haenszel statistic (Q_{CMH}) (8,9). The procedure (8) consisted of the following steps.

1. For each candidate independent variable (V), a Pearson chi-square was computed to test the hypothesis of independence between V and accident severity by using

$$X_p^2 = \Sigma[(\text{Observed} - \text{expected})^2/\text{Expected}] \quad (1)$$

The first variable selected was the one that had the largest X_p^2 per degree of freedom.

2. For each of the remaining variables (U), two-way tables were formed between variable U and accident severity for all levels of the first variable selected. The statistic Q_T was then computed by using

$$Q_T = \sum_{h=1}^{q} \underset{\sim}{G}_h' [\underset{\smile}{\mathrm{Var}}(\underset{\sim}{G}_h | H_0)]^{-1} \underset{\sim}{G}_h \quad (2)$$

where $\underset{\sim}{G}_h$ represents a matrix of the differences between the observed and the expected frequencies under H_0 for the hth of q tables, and $\underset{\sim}{G}_h'$ is the transpose matrix of $\underset{\sim}{G}_h$. The null hypothesis for Q_T states that, for each of the separate levels of the covariable set h = 1, 2, ..., q, the response variable is distributed at random with respect to the factor variables. The variable selected was the one with the largest value of Q_T. In this context Q_T reflects both the main effect of a specific variable and its interaction with the previously selected variables. However, after the first few variables were selected, Q_T might lose its asymptotic chi-square property because of the rapidly increasing degrees of freedom and the thinning out of the data in the cells of the contingency table. When this happened, Q_{CMH} could be used instead because it is not affected by the thinning process. It reflects the average partial association effect of a variable, as opposed to Q_T's total contribution, which includes interactions with other variables. Q_{CMH} is expressed as

$$Q_{CMH} = \underset{\sim}{G}' [\underset{\smile}{\mathrm{Var}}(\underset{\sim}{G} | H_0)]^{-1} \underset{\sim}{G} \quad (3a)$$

where

$$\underset{\sim}{G} = \sum_{h=1}^{q} \underset{\sim}{G}_h \quad (3b)$$

When using Q_{CMH} as the selection criterion, the variable exhibiting the largest value of Q_{CMH} was chosen.

3. The procedure given in 2 was repeated until the list of candidate variables was exhausted.

Models for Accident Severity

The model form selected for the severity analysis was a pair of logit models for a polytomous severity variable whose categories (fatal, injury, and property damage) have a natural order. Continuous ratios, such as fatal accidents to all nonfatal accidents and nonfatal injury accidents to property-damage accidents, have been shown by Fienberg (10) to possess the asymptotic properties that allow two logit models to be estimated independently while retaining the basic structure of the entire data set. The two logits were (a) that for the probability of fatal accident involvements and (b) that for the probability of nonfatal injury accident involvements. This model form was selected because, although fatal accidents are the ones for which the best understanding is most desired, they account for only a small proportion of total accidents. A single estimated severity model will tend to predict severe accidents far less well than nonsevere accidents.

In a recent study by Gimotty and Chirachavala (11), an estimated logit model for accident severity of passenger-car occupants almost always correctly predicted nonsevere injuries, but it correctly predicted severe injuries less than 50 percent of the time. Another advantage of the continuous ratios is that the independently estimated logits can be relatively more simple than the single severity model. This is desirable because it is easier to interpret a simple model (10). The models for accident severity can be represented by

P(fatal accident | an accident) = f(x_1, x_2, ...),
and
P(nonfatal injury accident | a nonfatal accident) = f(x_1, x_2, ...).

In these calculations x_1, x_2, ... are the independent variables.

The logit models for a dependent variable (whose index i = 1, 2, and 3 designates fatal, injury, and property-damage accident involvements, respectively) can be represented by

$$\mathrm{Log}[m_{1jk}/(m_{2jk} + m_{3jk})] = W + W_{2(j)} + W_{3(k)} + W_{23(jk)} \quad (4)$$

$$\mathrm{Log}(m_{2jk}/m_{3jk}) = W + W_{2(j)} + W_{3(k)} + W_{23(jk)} \quad (5)$$

where

- j and k = index levels of variables 2 and 3;
- m_{ijk} = expected cell frequency;
- $m_{1jk}/(m_{2jk} + m_{3jk})$ = ratio of fatal to nonfatal accident involvements; it is called the fatality ratio;
- m_{2jk}/m_{3jk} = ratio of nonfatal injury to property-damage accident involvements; it is called the injury ratio;
- W = overall mean;
- $W_{2(j)}$ = main effect of the jth level of variable 2;
- $W_{3(k)}$ = main effect of the kth level of variable 3; and
- $W_{23(jk)}$ = interaction effect between the jth level of variable 2 and the kth level of variable 3.

The models as represented by Equations 4 and 5 express the probability of a fatal accident involvement given an accident involvement and the probability of a nonfatal injury accident involvement when the involvement is nonfatal.

To ensure that the estimated logit models for the probability of a fatal accident involvement and the probability of an injury accident involvement are integral parts of the same contingency table, the following modeling estimation procedure was adopted. First, a log-linear model for a full contingency table was estimated to determine the intrinsic associations among all the independent variables. These associations were then included in estimating the logit models for the two probabilities. In estimating the models, the total sample size of the contingency table was assumed fixed.

VARIABLE SELECTION

The 18 independent variables initially considered in the analysis are given in Table 1. They represent accident, vehicle, operational, driver, road, and environmental factors. Accident severity--the dependent variable--was defined to have three levels, as follows:

1. Accident involvements that result in at least a fatality,
2. Accident involvements that result in at least one injury but no fatality, and
3. Accident involvements that result in only property damage that exceeds $2,000.

TABLE 1 Candidate Independent Variables

Variable	Level
Accident: accident type	Single-vehicle, collision with car, collision with commercial vehicle
Vehicle	
Configuration	Straight truck
Trailer body style	Semitrailer, double van, flatbed, tanker, other
Gross vehicle weight	
Load status	Empty, loaded
Operation	
Trip length	Over-the-road, local
Cargo type	General cargo, solid bulk, liquid bulk, metal, chemicals, other
Driver	
Age	
Years of employment	
Hours on duty before accident	
Scheduled time of driving	
Road	
Road class	1-3 lanes or undivided, 4 or more lanes, divided
Road surface condition	Dry, wet, snow
Accident on ramp	Yes, no
Environment	
Rural or urban	
Light condition	Day, night
Time of day	
Weather	Clear, overcast, rain, snow

Ten of the 18 independent variables considered were found to be significant by the variable selection process: accident type, vehicle configuration, trailer body style, gross vehicle weight, cargo weight, trip length, road class, road surface condition, rural or urban environment, and day or night.

The variables that were eliminated (at the 0.05 significant level) were driver age, scheduled time of driving, time of accident, weather, and ramp accident. Driver experience with the company and driver hours on duty were not significant at the 0.01 level and were also eliminated. Cargo type was excluded from further analysis because it was found to be highly correlated with vehicle configuration and trailer body style.

The problems of empty cells and the requirement of reasonable cell counts led to combining some of these 10 variables or the reduction of their number of levels or both. The six final independent variables that were used in the subsequent modeling and their univariate distributions are given in Table 2. The vehicle-type variable consists of four categories: large single-unit truck more than 10,000 lb, truck or tractor with one van-style trailer, truck or tractor with one flatbed- or tanker-style trailer, and tractor pulling two trailers. These four vehicle types will be referred to as straight truck, van single, flatbed or tanker single, and double, respectively. [Not shown in the table is the trip-length variable (local versus over-the-road operations).]

TABLE 2 Truck-Combination Accident Involvements by Variable

Variable	Level	No. of Involvements
Accident type (V2)	Single vehicle	9,891
	With car or motorcycle	10,398
	With commercial vehicle	4,031
Environment (V3)	Day and dry surface	9,389
	Night and dry surface	7,577
	Wet or snowy surface	7,891
Road class (V4)	Rural, 1-3 lanes or undivided	7,925
	Rural, 4 or more lanes divided	8,400
	Urban streets	8,967
Load status (V5)	Empty	7,108
	Loaded	18,888
Vehicle type	Straight truck	774
	Van Semitrailer	13,566
	Flatbed or tanker semitrailer	6,959
	Double	982

Note: Data are from BMCS, 1980.

MODEL ESTIMATION

Although a logit model that has seven categorical variables (six independent variables and one dependent variable) can be estimated, partitioning the data can yield a better understanding of the modeling results. High-order interaction effects in a contingency-table analysis are usually difficult to interpret and typically occur in the situation where there are many variables or the data are unevenly distributed among the cells or both (12).

The first partition of the data was by trip length (local versus over-the-road operations). These two types of operations differ significantly in the environment and operational aspects that may give rise to different severity patterns. Separate analyses were therefore carried out for local and over-the-road operations. Only the findings of over-the-road operation are reported in this paper. The next partition was by vehicle type because, when it was included as one of the variables in the model, a saturated model that exactly duplicated the contingency table was obtained.

The model terms that were found to be significant for the four truck types for the two severity measures are given in Table 3. In most models the main effects of accident type, environment, road class, the interaction between accident type and road class, and the interaction between road class and environment formed the basic common variable effects. Detailed discussion of the significant main effects and interactions is presented for each vehicle type in the following sections. The main effects of the significant variables are highlighted as follows.

The main effect of accident type on both fatality and injury ratios was significant in all eight models (the only such effect). Of the three accident

TABLE 3 Variables and Interactions in Severity Models

FATALITY RATIO MODELS

Vehicle Type	V2	V3	V4	V5	V2xV3	V2xV4	V2xV5	V3xV4	V2xV3xV4
Straight Truck	●	●	●						
Van Semi-Trailer	●	●	●			●		●	
Flat bed/Tanker Semi-Trailer	●	●	●	●	●	●	●	●	●
Double	●		●			●			

INJURY RATIO MODELS

Vehicle Type	V2	V3	V4	V5	V2xV3	V2xV4	V2xV5	V3xV4	V2xV3xV4
Straight Truck	●								
Van Semi-Trailer	●	●	●			●		●	
Flat bed/Tanker Semi-Trailer	●	●	●			●			
Double	●	●	●			●		●	

V2 - Accident Type
V3 - Environment
V4 - Road Class
V5 - Load Status

types, collisions with passenger cars always resulted in the highest fatality and injury ratios. The main effect of environment was significant in six of the eight models. Of the three environmental conditions, dry road surfaces at night usually had the highest fatality and injury ratios. The main effect of road class was significant in seven of the eight models. Of the three road classes, undivided rural roads usually had the highest fatality ratio for all truck types, followed by divided rural roads and urban streets. Hedlund (7) reported a similar finding on the effect of road class. For injury ratios, there was little difference between undivided and divided rural roads. Injury ratios for car-truck collisions on urban streets were as high as those on rural roads.

The interaction between road class and accident type was significant for all six combination-vehicle models. For single-vehicle accidents, there was little difference in the fatality ratios on different road classes, whereas multiple-vehicle collisions had much higher ratios on undivided rural roads. The interaction between road class and environment was significant in four out of six combination-vehicle models. Fatality ratios for wet/snowy pavements were usually lower on divided rural roads than would be expected; dry-surface urban streets during the day also had lower-than-expected fatality ratios.

Severity for Straight-Truck Accidents

There were 640 straight-truck involvements with complete information available on all variables. The models for straight trucks were the simplest of all. The fatality ratios for straight trucks are best explained by the independent main effects of accident type, road class, and environment. The data in Table 4 and Figure 1 give the fitted fatality ratios, which ranged from 0.003 to 0.67. The effect of accident type was such that fatality ratios for single-truck accidents, collisions with a commercial vehicle, and collisions with a car were 1:4:19. The effect of environment was such that fatality ratios for day/dry, wet/snowy, and night/dry were 1:1.5:2.5. The effect of road class was such that the ratios for urban, rural divided, and rural undivided roads were 1:1:5. Note that in this data set two out of five collisions between straight trucks and cars that occur on undivided rural roads at night on dry pavement can be expected to result in a fatality.

The fitted injury-ratio model for straight trucks depends only on the accident type (Table 4). The fitted injury ratio for with-car collisions was more than twice that for either single-vehicle accidents or for with-commercial-vehicle collisions.

Severity for Van Semitrailers

There were 11,748 accident involvements of van singles with complete information on all independent variables. Both the estimated fatality-ratio and injury-ratio models included the interaction between accident type and road class, the interaction between road class and environment, and the main effects of these three variables. The data in Table 5 give the fitted fatality ratios, and Figure 2 shows these effects. Fatality ratios for this vehicle type ranged from 0.01 to 0.26.

The interaction effect between accident type and road class on fatality ratios was such that, for single-vehicle accidents, the ratios for urban, divided rural, and undivided rural roads had a range from 1:1:1 to 0.7:1.6:1, depending on the particular environmental condition. For collisions with cars or with commercial vehicles, the differences in the ratios among these road classes were greater--from 0.4:0.4:1 to 0.3:1:1. The fatality ratios were

TABLE 4 Estimated Fatality and Injury Ratios for Straight Trucks

Road	Environment	Accid. Type	Fatality Ratio	Injury Ratio
Rural Undivided	Dry Day	Single Veh.	0.014	1.13
		With Car	0.259	2.69
		With Comm.	0.054	1.10
	Dry Night	Single Veh.	0.036	*
		With Car	0.667	
		With Comm.	0.139	
	Wet/Snowy	Single veh.	0.023	*
		With Car	0.420	
		With Comm.	0.088	
Rural Divided	Dry Day	Single Veh.	0.003	*
		With Car	0.055	
		With Comm.	0.012	
	Dry Night	Single Veh.	0.008	*
		With Car	0.143	
		With Comm.	0.030	
	Wet/Snowy	Single Veh.	0.005	*
		With Car	0.090	
		With Comm.	0.018	
Urban	Dry Day	Single Veh.	0.003	*
		With Car	0.061	
		With Comm.	0.013	
	Dry Night	Single Veh.	0.008	*
		With Car	0.158	
		With Comm.	0.033	
	Wet/Snowy	Single Veh.	0.005	*
		With Car	0.099	
		With Comm.	0.021	

*Same as entry above

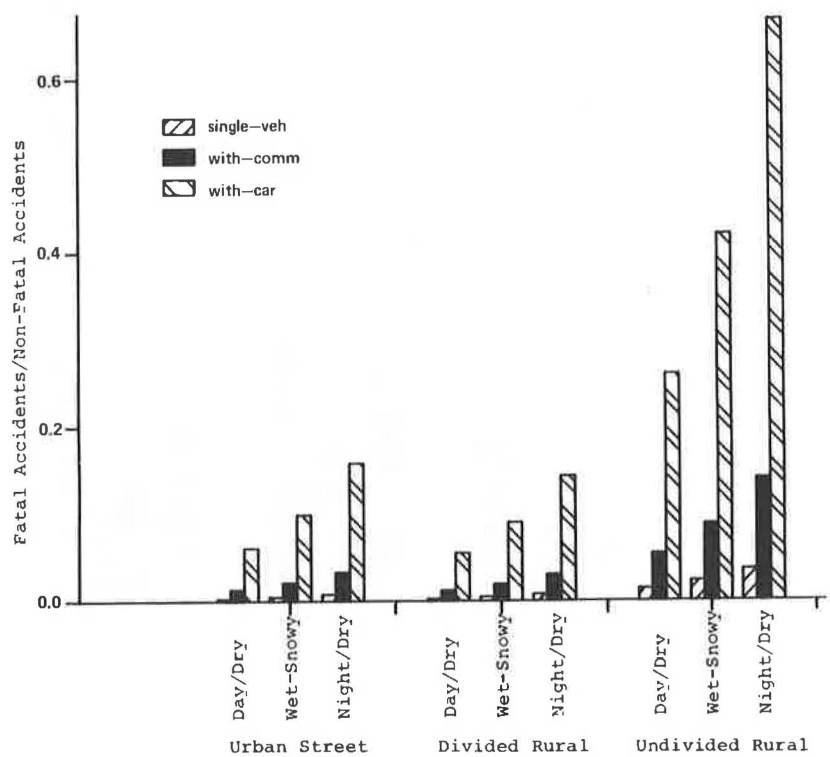

FIGURE 1 Estimated fatality ratios for straight trucks.

TABLE 5 Estimated Fatality and Injury Ratios for Van Semitrailers

Road	Environment	Accid. Type	Fatality Ratio	Injury Ratio
Rural Undivided	Dry Day	Single Veh.	0.023	1.01
		With Car	0.150	2.55
		With Comm.	0.089	1.37
	Dry Night	Single Veh.	0.040	1.31
		With Car	0.261	3.29
		With Comm.	0.155	1.76
	Wet/Snowy	Single veh.	0.022	0.98
		With Car	0.147	2.46
		With Comm.	0.087	1.32
Rural Divided	Dry Day	Single Veh.	0.034	1.14
		With Car	0.102	2.54
		With Comm.	0.079	1.36
	Dry Night	Single Veh.	0.049	1.28
		With Car	0.148	2.86
		With Comm.	0.115	1.54
	Wet/Snowy	Single Veh.	0.021	0.80
		With Car	0.064	1.77
		With Comm.	0.050	0.95
Urban	Dry Day	Single Veh.	0.014	0.50
		With Car	0.035	2.63
		With Comm.	0.022	0.63
	Dry Night	Single Veh.	0.042	0.54
		With Car	0.106	2.86
		With Comm.	0.065	0.69
	Wet/Snowy	Single Veh.	0.019	0.55
		With Car	0.049	2.91
		With Comm.	0.030	0.70

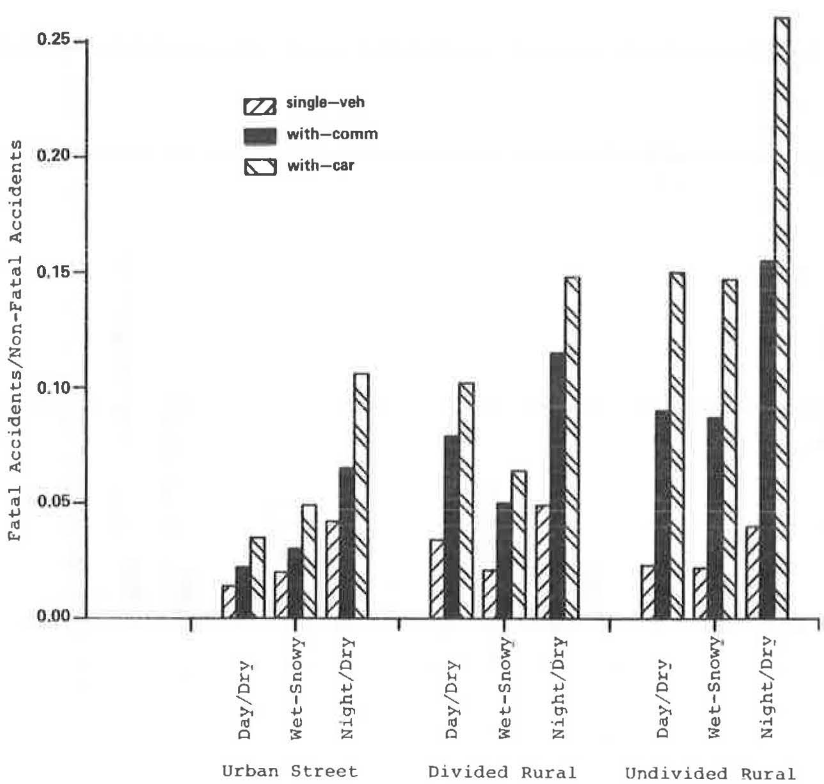

FIGURE 2 Estimated fatality ratios for van singles.

usually the highest for undivided rural roads, followed by divided rural and urban roads.

The interaction effect between environment and road class was such that, for all road classes, there was a much higher fatality ratio for dry pavements at night than for the other two conditions. This is particularly true in two-vehicle collisions. Day accidents on dry pavements had a higher ratio than did the wet/snowy condition for divided rural roads. For undivided rural and urban roads, the difference between these two environmental conditions was small. Of note is a relatively high fatality ratio for collisions with other commercial vehicles on all rural roads at night (0.16 for undivided rural and 0.12 for divided rural roads). These are rare contrasts in which the with-car fatality ratio is not completely dominant for a road class.

The fitted injury ratios for van singles are given in Table 5 and shown in Figure 3. The interaction effect between environment and road class on injury ratios was such that, on all rural roads, there was usually a higher injury ratio at night than during the day on dry pavements; the latter, in turn, was much higher than under wet/snowy condition. For urban roads, all three environmental conditions had similar injury ratios.

The interaction effect between accident type and road class on injury ratios was that, for single-vehicle accidents, the injury ratios on urban, divided rural, and undivided rural roads ranged from 0.5:0.8:1 to 0.5:1:1. For with-car collisions they ranged from 1.1:0.7:1 to 1:1:1. For with-commercial-vehicle collisions they ranged from 0.5:0.7:1 to 0.5:1:1.

Severity for Flatbed or Tanker Semitrailers

There were 6,041 accident involvements of flatbed or tanker semitrailers with complete information on all variables. The fatality ratios are best explained by a two-factor interaction between accident type and load status and a three-factor interaction involving accident type, environment, and road class. The fitted fatality ratios are given in Table 6.

The interaction effect between accident type and load status on fatality ratios was that for single-vehicle and with-commercial-vehicle accidents. Loaded flatbed or tanker singles had twice as high a fatality ratio as did the empty singles, whereas the fatality ratio for with-car collisions did not change much for loaded vehicles. Figure 4 shows this interaction effect on undivided rural roads. Loaded flatbed or tanker singles also had relatively high fatality ratios for with-commercial-vehicle collisions on all rural roads at night on dry pavements.

The three-factor interaction effect among accident type, environment, and road class on fatality ratios was such that dry pavements at night had a higher fatality ratio than did other environmental conditions for all combinations of accident type and road class. The only exception was with-car collisions on undivided rural roads, where there was little difference in fatality ratios for the different environmental conditions.

The injury-ratio model for flatbed or tanker singles included the main effects of accident type, road class, and environment together with the interaction between accident type and road class. The data in Table 6 and in Figure 5 give the fitted injury ratios.

The main effect of environment on injury ratios was that the night/dry condition had higher injury ratios than did the other environmental conditions. The interaction effect between accident type and road class was that single-vehicle accident severity on both types of rural roads was almost 50 percent higher than that found on urban roads. With-car collisions on undivided roads had a similar ratio to

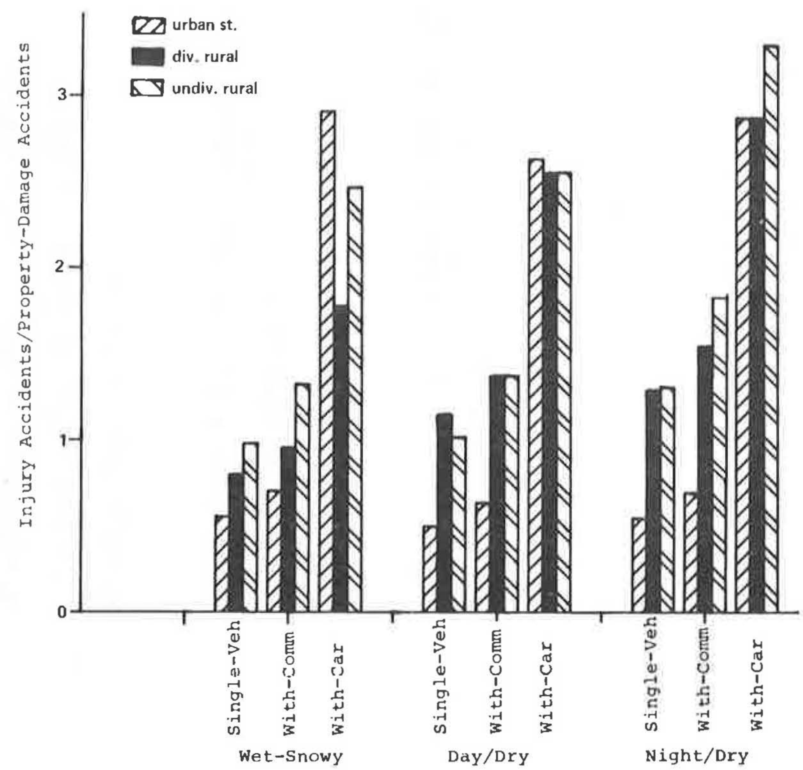

FIGURE 3 Estimated injury ratios for van singles.

TABLE 6 Estimated Fatality and Injury Ratios for Flatbed or Tanker Semitrailers

Road	Environment	Accid. Type	Fatality Ratio		Injury Ratio
			Empty	Loaded	
Rural Undivided	Dry Day	Single Veh.	0.037	0.072	0.78
		With Car	0.153	0.170	2.31
		With Comm.	0.046	0.095	0.96
	Dry Night	Single Veh.	0.047	0.094	1.03
		With Car	0.150	0.166	3.05
		With Comm.	0.069	0.143	1.28
	Wet/Snowy	Single Veh.	0.008	0.016	0.67
		With Car	0.155	0.172	1.98
		With Comm.	0.035	0.071	0.83
Rural Divided	Dry Day	Single Veh.	0.029	0.058	0.74
		With Car	0.086	0.096	2.22
		With Comm.	0.033	0.068	1.23
	Dry Night	Single Veh.	0.039	0.078	0.98
		With Car	0.142	0.158	2.93
		With Comm.	0.069	0.142	1.63
	Wet/Snowy	Single Veh.	0.005	0.011	0.63
		With Car	0.046	0.051	1.90
		With Comm.	0.043	0.087	1.06
Urban	Dry Day	Single Veh.	0.002	0.004	0.51
		With Car	0.054	0.060	2.23
		With Comm.	0.026	0.054	0.51
	Dry Night	Single Veh.	0.023	0.046	0.67
		With Car	0.103	0.114	2.96
		With Comm.	0.028	0.058	0.67
	Wet/Snowy	Single Veh.	0.016	0.033	0.44
		With Car	0.049	0.054	1.92
		With Comm.	0.000	0.000	0.44

those on other roads. The injury ratios for with-commercial-vehicle accidents on urban, divided rural, and undivided rural roads were in a 0.5:1.2:1 proportion.

Severity for Doubles

There were 834 accident involvements of doubles with complete information on all variables. The fatality ratios are best explained by the main effects of accident type and road class, and their interaction. The data in Table 7 and in Figure 6 give the fitted fatality ratios, which ranged in values up to 0.22.

The interaction effect between accident type and road class on fatality ratios was that, for single-vehicle accidents, the ratios on urban, divided rural, and undivided rural roads were about 2:1:1. For with-car collisions they were about 0.2:0.9:1. For with-commercial-vehicle collisions they were 0:0.6:1.

The injury ratio for doubles is best explained by two complex interactions involving road class with accident type and environment. The data in Table 7 and in Figure 7 give the fitted injury ratios. Of interest is that the injury ratios for two-vehicle collisions on wet/snowy urban streets were extremely high (4.8 for with-car and 2.6 for with-commercial-vehicle collisions). Two-vehicle collisions on dry-surfaced divided rural roads at night had considerably higher injury ratios than did those on undivided rural roads. Doubles are the only truck type of which with-car collisions on divided rural roads had a considerably higher injury ratio than did those on undivided rural roads.

SUMMARY AND DISCUSSION OF RESULTS

Of the 18 candidate accident, driver, vehicle, road,

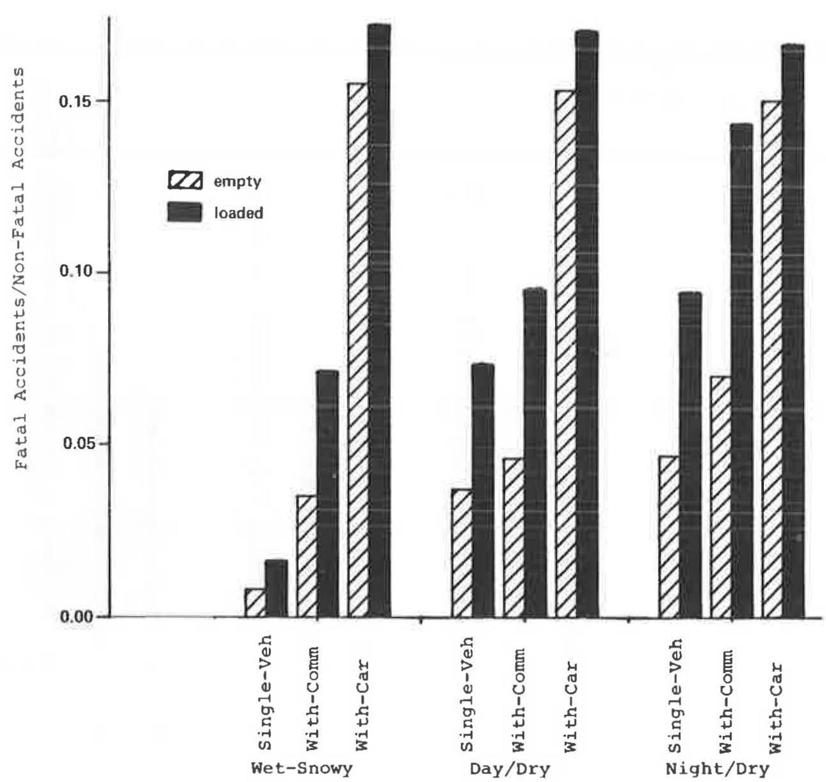

FIGURE 4 Estimated fatality ratios for flatbed or tanker singles.

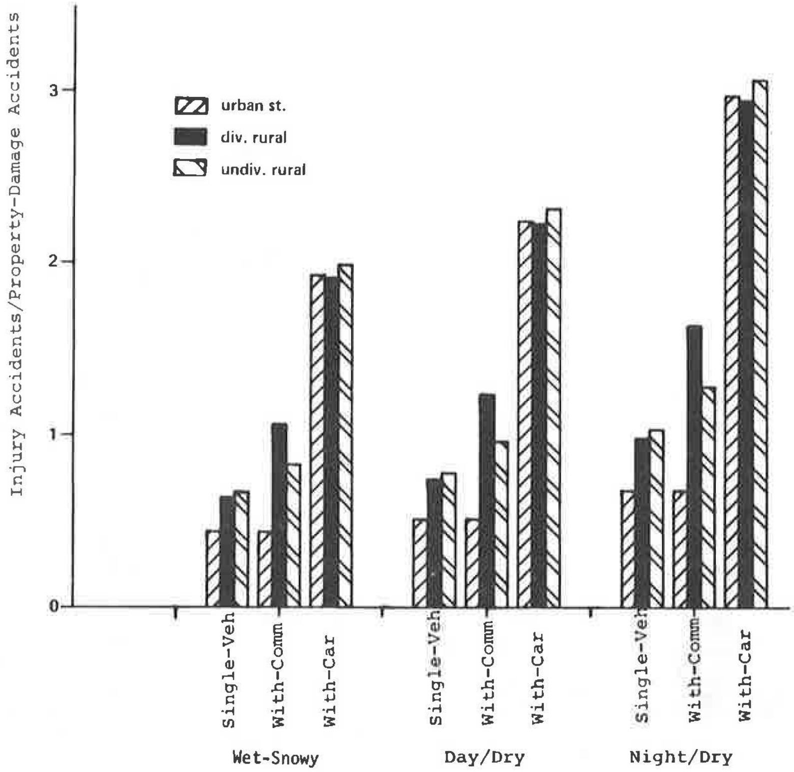

FIGURE 5 Estimated injury ratios for flatbed or tanker singles.

TABLE 7 Estimated Fatality and Injury Ratios for Doubles

Road	Environment	Accid. Type	Fatality Ratio	Injury Ratio
Rural Undivided	Dry Day	Single Veh. With Car With Comm.	0.031 0.222 0.111	0.75 2.35 1.24
	Dry Night	Single Veh. With Car With Comm.	*	0.74 2.32 1.23
	Wet/Snowy	Single veh. With Car With Comm.	*	0.78 2.44 1.30
Rural Divided	Dry Day	Single Veh. With Car With Comm.	0.042 0.186 0.065	0.48 1.28 0.71
	Dry Night	Single Veh. With Car With Comm.	*	1.44 3.79 2.10
	Wet/Snowy	Single Veh. With Car With Comm.	*	0.98 2.57 1.43
Urban	Dry Day	Single Veh. With Car With Comm.	0.080 0.036 0.000	0.27 2.01 1.06
	Dry Night	Single Veh. With Car With Comm.	*	0.28 2.12 1.13
	Wet/Snowy	Single Veh. With Car With Comm.	*	0.64 4.77 2.55

*Same as entry above

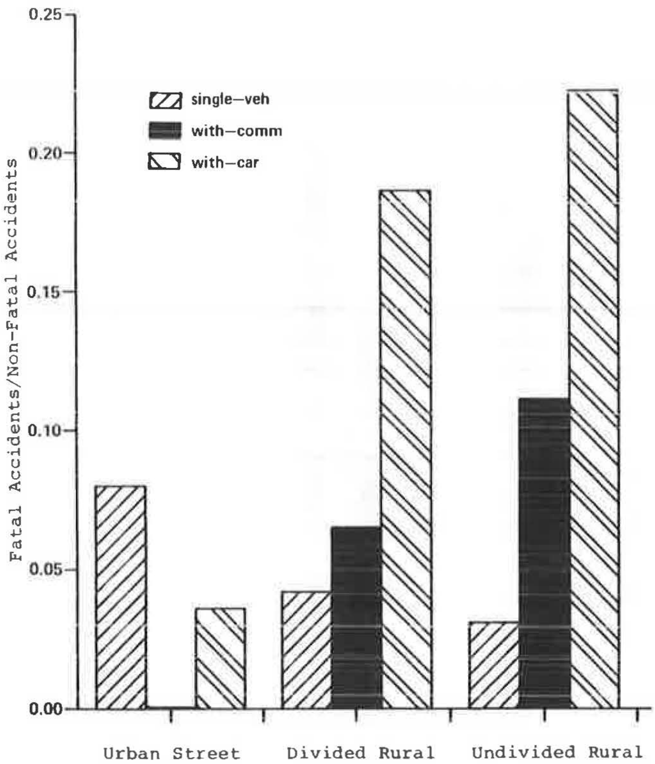

FIGURE 6 Estimated fatality ratios for doubles.

and environment variables considered, those significantly influencing the accident severity of large trucks and combination vehicles engaged in over-the-road operation were vehicle type, accident type, environment, road class, and, to a limited extent, load status. On the other hand, specific driver variables such as age, experience, hours of driving before accident, and scheduled driving time, as well as weather and time-of-day, were not found to be significant. The finding on the effect of road class is generally consistent with that reported by Hedlund (7). The effect of vehicle weight, which had been reported by Hedlund to be nonsignificant, has been found here to be significant for flatbed or tanker singles.

This analysis revealed that relatively simple and similar model structures were effective in capturing the severity differences among the accidents involving most types of large trucks and combination vehicles. The models for the different truck types are easy to interpret, and the differences are large enough to be of practical interest to decision makers.

The models for all combination vehicles, except flatbed or tanker singles, share a basic model form--the main effects of type of collision, road class, environment, and the interactions involving the road class. These effects alone explain most of the variation in accident severity.

The models for straight trucks have no interactions. They include only the main effects of the same three variables: accident type, road class, and environment. The three effects were that fatality ratios for straight trucks were much higher for collisions with cars, for accidents on undivided rural roads, and for accidents on dry pavements at night. The injury ratio for straight trucks was only affected by accident type. With-car collisions had about twice as high an injury ratio as did the other two types of accidents.

Severity models for van singles and doubles shared many common factors--the main effects of accident type, road class, and environment; the interaction between accident type and road class; and the interaction between road class and environment. Fatality ratios of van singles in two-vehicle collisions were higher on undivided rural roads than the

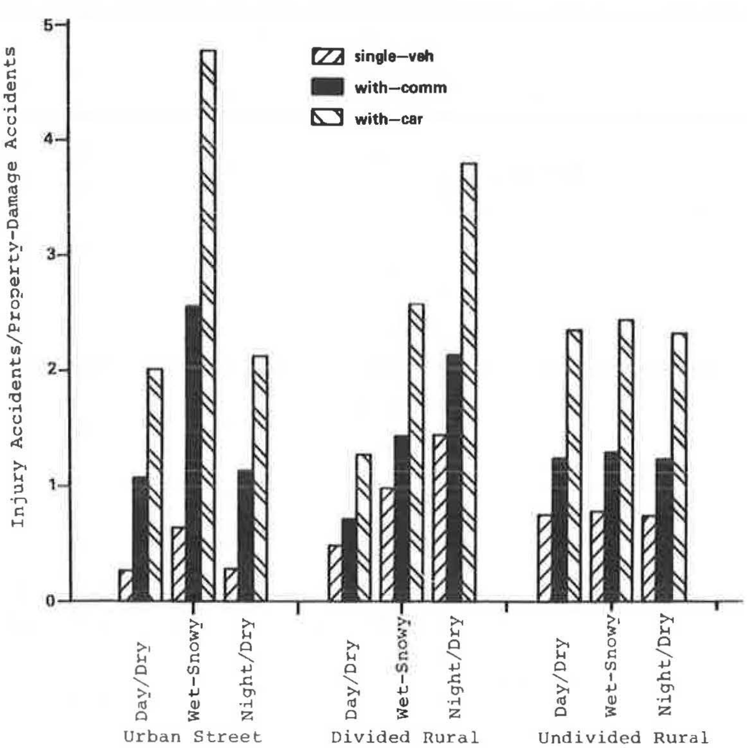

FIGURE 7 Estimated injury ratios for doubles.

other road classes. These collisions also had much higher fatality and injury ratios at night on dry pavements than under the other conditions. Two-vehicle collisions involving at least one double on both types of rural roads had similar fatality ratios. They were about 5 times as high as those found on urban streets. This type of collision on divided rural roads had a higher injury ratio at night on dry pavements than under the other conditions.

The severity models for flatbed or tanker singles were the most complex. In addition to the same three main effects and two interactions for van singles and doubles, the main effect of load status, the interaction between accident type and environment, the interaction between accident type and load status, and the three-way interaction involving accident type, road class, and environment were also significant for the fatality-ratio model for more than 6,000 flatbed or tanker single-accident involvements.

The accident severity for all vehicle types is summarized in Table 8 for the combinations of the other variables. The first symbol in each cell indicates the rating for the fatality ratio, and the second symbol indicates the rating for the injury ratio. The ratings of low, medium, high, and dangerous are based on the following scales of fatality and injury ratios:

1. For fatality ratios: 0 to 0.05 = low (L), 0.06 to 0.15 = medium (M), 0.16 to 0.25 = high (H), and 0.26 to 0.70 = dangerous (D); and
2. For injury ratios: 0 to 1.0 = low (L), 1.1 to 2.0 = medium (M), 2.1 to 3.0 = high (H), and 3.1 to 4.8 = dangerous (D).

The data in Table 8 permit easy identification of the accident characteristics that were most severe:

1. Straight-truck and passenger-car collisions on undivided rural roads under all environmental conditions,
2. Van single and passenger-car collisions on undivided rural roads at night,
3. Double and passenger-car collisions on undivided rural roads under all environmental conditions,
4. Double and passenger-car collisions on divided rural roads under all environmental conditions, and
5. Loaded flatbed or tanker single and passenger-car collisions on undivided rural roads under all environmental conditions.

The vehicle type with the lowest severity record was the empty flatbed or tanker singles. At the other end of the scale were doubles and straight trucks.

For all vehicle types, with-car collisions were the most severe, followed generally by with-commercial-vehicle collisions and single-vehicle accidents. This is because of the vulnerability of occupants of cars when colliding with these large vehicles. This is not surprising because the difference between the weight of a passenger car and the smallest BMCS truck is substantial. With-car collisions on undivided rural roads were usually more severe than those on divided rural roads or urban streets. On undivided rural roads collisions between passenger cars and straight trucks, doubles, and loaded flatbed or tanker singles were always highly severe under all conditions. This is also true for collisions on undivided rural roads at night between passenger cars and van singles. Many of these accidents are head-on. On divided rural roads collisions between passenger cars and doubles were very severe, more so than collisions between cars and other truck types. On urban streets collisions between passenger cars and any truck type resulted in a medium severity level. Prevailing speeds are probably the moderating factor in the severity of these collisions.

Collisions involving commercial vehicles and doubles, loaded flatbed and tanker singles, or van singles on rural roads were usually more severe than those involving commercial vehicles and straight

TABLE 8 Rating of Fatality and Injury Rates for Over-the-Road Operation

			Straight Trucks		Van Semitrailers		Flatbed or Tanker Semitrailers			Doubles	
							Fatality				
Accident Type	Road	Environment	Fatality	Injury	Fatality	Injury	Empty	Loaded	Injury	Fatality	Injury
Single vehicle	Rural divided	Dry day	L	M	L	M	L	M	L	L	L
		Dry night	L	M	L	M	L	M	L	L	M
		Wet/snowy	L	M	L	L	L	L	L	L	L
	Rural undivided	Dry day	L	M	L	L	L	M	L	L	L
		Dry night	L	M	L	M	L	M	L	L	L
		Wet/snowy	L	M	L	L	L	L	L	L	L
	Urban streets	Dry day	L	M	L	L	L	L	L	M	L
		Dry night	L	M	L	L	L	L	L	M	L
		Wet/snowy	L	M	L	L	L	L	L	M	L
With car	Rural divided	Dry day	L	H	M	H	M	M	H	H	M
		Dry night	M	H	M	H	M	M	H	H	D
		Wet/snowy	M	H	M	M	L	L	M	H	H
	Rural undivided	Dry day	D	H	M	H	M	H	H	H	H
		Dry night	D	H	H	D	M	H	H	H	H
		Wet/snowy	D	H	M	H	M	H	M	H	H
	Urban streets	Dry day	L	H	L	H	L	L	H	L	L
		Dry night	M	H	M	H	M	M	H	L	M
		Wet/snowy	M	H	L	H	L	L	M	L	D
With commercial vehicle	Rural divided	Dry day	L	M	M	M	L	M	M	M	L
		Dry night	L	M	M	M	M	M	M	M	H
		Wet/snowy	L	M	L	L	L	M	M	M	M
	Rural undivided	Dry day	L	M	M	M	L	M	L	M	M
		Dry night	M	M	M	M	M	M	M	M	M
		Wet/snowy	M	M	M	M	L	M	L	M	M
	Urban streets	Dry day	L	M	L	L	L	L	L	L	M
		Dry night	L	M	L	L	L	L	L	L	M
		Wet/snowy	L	M	L	L	L	L	L	L	H

Note: L = low, M = medium, H = high, and D = dangerous.

trucks or empty tanker or flatbed singles. However, on urban streets such collisions were more severe when involving doubles or straight trucks than other truck types.

For single-vehicle accidents on rural roads, loaded tanker or flatbed singles had a relatively higher severity level than other truck types. On urban roads doubles had a higher severity level than other truck types.

Note that injury and fatality ratios for accidents on urban streets may conceivably be sensitive to the BMCS reporting process. If the noninjury accidents on urban streets often resulted in property damage less than $2,000, the reported urban-street accidents would have a disproportionately larger fraction of injury and fatality accidents than would the rural-road accidents. This will inflate the injury and fatality ratios for urban streets relative to rural roads. This possible bias, however, will not affect any comparison of severity of the accidents on urban streets made across the other independent variables.

CONCLUSIONS

The need for measures to reduce the severity of car-truck collisions is particularly important for undivided rural roads and urban streets. Such collisions on undivided rural roads usually result in a high incidence of fatality or bodily injuries to car occupants. Although car-truck collisions on urban streets are not usually as severe as those on rural roads, the large number of passenger cars on urban streets can result in a large number of accident casualties.

On undivided rural roads collisions between passenger cars and straight trucks, doubles, and loaded tanker or flatbed singles are highly severe under all conditions, particularly at night. Collisions between passenger cars and van singles are also highly severe at night on undivided rural roads. These point to a need for improved safety measures for nighttime operation of these large vehicles and passenger cars on these roads.

The car-truck collisions on divided rural roads are relatively less serious than those on undivided rural roads, with the exception of collisions between cars and doubles, which have a high severity level. Because doubles are expected to increase in number in the future, special consideration should be given to countermeasures to reduce both the number of such collisions and their severity.

The complex model for flatbed or tanker singles indicates that the severity of the accidents involving this class of vehicles is different from other truck types and warrants further investigation. In particular, attention to measures that will reduce the severity of accidents involving loaded flatbed or tanker singles deserve special consideration because of their higher severity level. Not only is this high severity rate a cause for concern, but because these vehicles are usually engaged in carrying fuels, solid or liquid bulk, logs and poles, metal, and other heavy products, accidents involving these vehicles carry the extra risk of cargo spillage, with possible dangerous effects to the land and communities near the accident site, as well as to other traffic.

ACKNOWLEDGMENT

The authors wish to thank James O'Day of the Transportation Research Institute, University of Michigan, for sharing his insight and knowledge of large-truck accidents with the authors.

Discussion

Olga J. Pendleton*

Chirachavala, Cleveland, and Kostyniuk are to be commended on using a relatively new but powerful and effective statistical technique for analyzing accident severity data. One of the factors that has stifled more extensive use of this method is the availability of computer software for its implementation. It would be informative if the authors would reveal their source of computer software for doing this.

The only points I would like to raise in this discussion involve the variable subset selection process. They are more of an academic nature and are presented as material for future research.

The variable selection procedure described in this study resembles the forward selection procedure in least-squares regression. It is natural to ask, then, if these procedures share the same problems. It is well-known that the forward selection procedure, as applied to regression analysis, may not reveal the best subset because it considers the interrelationship of variables in a sequential manner. That is, if a variable (for example, X_5) is selected as most important according to the the χ^2 test of step 1, it is fixed in the model, and only combinations of that variable and all others are considered as candidate models in the selection process. Suppose a variable that was not significant in step 1 (for example, X_3) would be significant in combination with another variable (for example, X_1), and more significant perhaps than the combination of X_5 and any other variables. The combination of X_3 and X_1 would never be examined as a candidate model by using the current procedure.

The computational problem of examining all possible models is prohibitive both here and in regression. It was circumvented in regression, however, by the discovery of an algorithm that eliminated the necessity of examining all models, but still guaranteed the selection of the best subset based on the criteria of minimum mean squared error (13). Discovery of a similar algorithm for the categorical subset selection process is a fascinating topic for future statistical research.

A second point I would like to raise concerns the treatment of continuous variables in this selection process. Obviously, these variables had to be categorized to accommodate the variable selection criteria. Does this classification affect the variable selection process? For example, if a variable such as driver age was initially classified into equal 20-year age groups and found to be nonsignificant in the variable selection process, would a reclassification of driver age into <20, 20 to 25, 25 to 40, >40 alter the importance of this variable? If so, how should continuous variables be classified? Should a different criteria such as a t-test (or Poisson test) for comparing equality of mean age among severity levels be applied to continuous data? How could that be combined with the categorical selection criteria?

Finally, does the redefinition of categories affect the subset selection process for the discrete variables? In this study classification categories were redefined for the final variables selected to

*Operations Research and Systems Analysis Program, Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843

accommodate zero cells. If the subset selection process were redone by using these new or collapsed categories, would the conclusions of the subset selection process change?

The questions raised in this discussion are difficult to answer but potentially important to the conclusions based on this type of analysis. It is surely beyond the scope of this study; however, this study has provided considerable insight by using a new and powerful statistical tool. In so doing, it has also opened a Pandora's box for statisticians!

Authors' Closure

The discussant is to be thanked for providing a stimulating discussion on the important aspects of the variable selection procedure not explicitly addressed in the paper.

The source of computer software used for the variable selection procedure and the modeling are PARCAT [Landis et al. (9)] and ECTA (Center for Population Studies, University of Michigan). Both are available from the University of Michigan, Ann Arbor, Michigan.

It is true that the variable selection procedure, because of its sequential nature, may not yield the best subset (based on some criteria). However, it is also likely to be true that the selected subset would be close to the best (if it was not the best) because the variables not selected will not significantly contribute to the explanation of the dependent variable in the presence of the selected variables. Besides, the criteria used to evaluate the best statistically often vary with circumstances and individuals' values. Additional test statistics or steps added to those in the existing algorithm that can define and assure the selection of the best subset will certainly be valuable.

When the selection process involves a continuous variable, consideration has to be given to many pertinent questions: how predominant the effect of this variable is relative to the other variables, and whether the effect is truly continuous (i.e., the effect is sensitive to a relatively small change in the values of the variable) or more step-function like. Here the effects of the variables such as driver age, number of years of employment, and scheduled time of driving were more likely to be step-function in nature. Their classifications were aided by past studies, judgment of experts, and examination of the data. In general, the classification adopted for a continuous variable should also be directed at answering (or testing) the substantive questions of interest. It is believed that a t-test and a Poisson test here might not yield conclusive classifications because, at that stage, the effects of other confounding variables were not sufficiently known to permit relatively homogeneous subsets for effective t-tests or Poisson tests. After all, such knowledge was partly what the subsequent modeling task was trying to achieve. In a case where a particular continuous variable is believed to be overwhelmingly important, then an analysis such as that used in Gimotty and Chirachavala (11), which incorporates both continuous and categorical variables, may be considered.

Finally, the redefined categories to accommodate structured zero cells or small cells actually make the modeling results more reliable. This is so because the nature of a chi-square test requires that the expected frequency in each cell be of a reasonable size. The redefinition can conceivably alter the values of the test statistic, the degrees of freedom, and perhaps the P-value of the test if the selection process is redone. But, unless the reclassification is quite extreme or different from the original definition, the significance of a particular variable should not be altered. However, if the researcher is uncertain of the outcome, an analysis such as factor analysis can be used at this step to assist in the reclassification. It is always difficult to generalize the effect of reclassification of variables for all situations without knowing the extent of and the reasons for the reclassification as well as the nature of the variables and the study.

These comments are much more specific to the circumstances that prevailed in this paper. Generalized answers for general situations are not likely to be simple, as the discussant has already pointed out. Statistical tests or models, no matter how sophisticated, are only simplified tools that assist researchers in making inferences on the populations. As such, their applications should go hand in hand with the substantive knowledge in the area of study.

REFERENCES

1. T. Chirachavala. Analysis of Accidents Involving Large Trucks and Combination Vehicles: A Discrete Multivariate Analysis. Ph.D. dissertation. Department of Civil Engineering, University of Michigan, Ann Arbor, 1984.
2. R. Zeiszler. Accident Experience of Double Bottom Trucks in California. California Highway Patrol, Sacramento, April 1973.
3. A Safety Comparison of Doubles Versus Tractor-Semitrailer Operations. Tech. Report. Bureau of Motor Carriers Safety, FHWA, U.S. Department of Transportation, Nov. 1977.
4. S.Y. Chang, M.L. Reiss, and H.W. McGee. Comparison of California Accident Rates for Single and Double Tractor-Trailer Combination Trucks. Final Report. Biotechnology, Inc., Falls Church, Va., March 1978.
5. G.R. Vallette, H.W. McGee, J.H. Sanders, and D.J. Enger. The Effect of Truck Size and Weight on Accident Experience and Traffic Operation--Volume III: Final Report. Biotechnology, Inc., Falls Church, Va., 1982.
6. J.C. Glennon. Trial Record. In The Supreme Court of the United States: Appendix. No. 79-1320. On Appeal from the United States Court of Appeals for the Eighth Circuit, October Term, 1979.
7. J. Hedlund. The Analysis of Large Truck Accidents. Mathematical Analysis Division, National Center for Statistics and Analysis, Research and Development, NHTSA, U.S. Department of Transportation, April 1977.
8. J.E. Higgins and G.G. Koch. Variable Selection and Generalized Chi-Square Analysis of Categorical Data Applied to a Large Cross-Sectional Occupational Health Survey. International Statistical Review, Vol. 45, 1977, pp. 51-62.
9. J.R. Landis, E.R. Heyman, and G.G. Koch. Average Partial Association in Three-Way Contingency Tables: A Review and Discussion of Alternative Tests. International Statistical Review, Vol. 46, 1978, pp. 237-254.
10. S.E. Fienberg. The Analysis of Cross-Classified Categorical Data. MIT Press, Cambridge, Mass, 1977.

11. P.A. Gimotty and T. Chirachavala. Injury Models for the Prediction of Severe Accidents. In Proceedings of the 26th Conference of the American Association of Automotive Medicine, 1982, pp. 63-76.
12. Y.M.M. Bishop, S.E. Fienberg, and P.W. Holland. Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, Mass., 1975.
13. R.R. Hocking and R.N. Leslie. Selection of the Best Subset in Regression Analysis. Technometrics, Vol. 9, 1967, pp. 531-540.

Problem of Identifying Hazardous Locations Using Accident Data

EZRA HAUER and BHAGWANT N. PERSAUD

ABSTRACT

Most agencies with responsibility for extensive road systems use some variant of the rate-and-number method to identify hazardous locations or blackspots. Sites so identified are later examined in detail to diagnose deficiencies and to suggest remedial measures. In this paper the degree to which the rate-and-number method is successful in identifying the unidentified, and what proportion of the sites that are subjected to detailed examination are not deviant at all, is examined. The first part of the paper is devoted to the analysis and development of the mathematical machinery. In the second part the use of the analytical results is illustrated by application to two data sets--one dealing with highway ramps in Ontario and the other with California drivers. The main result of this research is the facility to examine the performance of various identification procedures on the basis of measures of performance that are easy to understand. Such an examination should lead to a realistic assessment of what can be attained when identification for treatment is made on the basis of past accident history.

In most agencies with jurisdiction over extensive road systems it is common practice to try and rectify so-called accident blackspots. Ordinarily a two-stage process is used. In the first stage the past accident history of all sites is reviewed to select a limited number of apparently dangerous locations for further examination. In the second stage the selected sites are studied in more detail, often in the field, in order to devise cost-effective remedial projects for some of the sites.

The two-stage process is required because detailed examination of all sites is impractical. It is hoped that the first stage of the process will act as a sieve. A good sieve is one that allows through all sites that do not require remedial action and retains all sites that do require detailed study. Conversely, an inefficient sieve is one that retains a large number of sites that do not need close scrutiny and allows most blackspots to pass through its holes and thus escape identification. The purpose of this paper is to examine the quality and performance of a commonly used sieve.

Most sieves in current use are a variant of the rate-and-number method. Sites that register an unusually high number of accidents during a specified period of time or an unusually high accident rate (accidents per vehicle kilometer) are selected for inspection. Accidents are often weighted according to their severity. The rationale for the rate-and-number method appears to be left unspecified in the literature. However, a plausible line of reasoning for its raison d'être goes as follows:

> If the accident history of a site is found to deviate from the norm for its class, there surely is some reason for it. If so, a responsible agency and its professionals should examine the cause for this deviation and, if a cost-effective remedy can be found, should remove the cause of substandard performance.

It should be evident that a sieve that screens sites on the basis of number of accidents, accidents per vehicle kilometer, or accident severity is aimed only at establishing deviancy. It is not an indication of how easy or how difficult remedial treatment might be.

Some causes of substandard performance are random and fleeting in nature and essentially unrelated to the physical characteristics of the site. (Consider, for example, a local snow squall or rainstorm that causes several accidents to occur within a few minutes.) Other causes for deviation from the norm are more permanent in nature (sharp curves, polished pavement, narrow bridge, and so forth). These are the causal factors that are subject to remedial action. Accordingly, the object of the exercise is to identify sites for which the deviation from the safety norm is attributable to some permanent properties of the site.

It is well-known that the actual number of accidents occurring on a site fluctuates from year to year. It is only the average number of accidents in the long run that can be linked to the permanent properties of the site. This gives rise to the fundamental difficulty facing the screening process.

> Researchers wish to identify those sites for which, say, the "average number of accidents in the long run" deviates from the norm. However, in the identification process, researchers are restricted to the use of accident histories that are subject to pronounced random fluctuation.

This inescapable difficulty affects the quality of all sieves. When the number of accidents occurring on a site in the last 2 or 3 years is higher than the average in the long run for that site, the site will be caught by the sieve and subjected to detailed inspection, possibly unnecessarily. Conversely, sites with permanent properties such that their average in the long run is considerably higher than the norm will often escape detection because of a random down-fluctuation.

Accordingly, several questions are raised: How good are the accident-history-based sieves for blackspot identification? Do they capture most of the truly deviant sites? How many normal sites are lumped with the deviant ones that are labeled blackspots? How many deviant sites escape detection?

These questions translate into the following figures of merit by which the quality of the sieve should be measured:

1. The number of sites selected for closer examination (this is a measure of the effort required at the later stage when site-specific deficiencies are identified, remedies are designed, and economics are examined);
2. The number of truly deviant sites among those selected for closer inspection (these are often called correct positives);
3. The number of sites that are not deviant yet have been captured by the sieve and selected for closer inspection (these are the false positives); and
4. The number of truly deviant sites that are not identified as requiring attention (these are called the false negatives).

The central issues are easy to visualize with the aid of a Venn diagram. Let the box in Figure 1 symbolize the collection of all sites. The set of all deviant sites is delimited by curve 1. Thus all nondeviant sites are outside curve 1. The set of all sites selected for closer inspection is enclosed by curve 2. An ideal sieve would be one for which curves 1 and 2 coincide. However, no real sieve or screening process is ideal. Therefore, curves 1 and 2 delineate three distinct sets. Set A contains sites that are deviant but are not selected for closer inspection. These are the false negatives. Set B contains deviant sites that are selected for inspection. These are the correct positives. Set C contains sites that are not deviant but are selected for inspection. These are the false positives.

The union of sets A and B is the collection of all deviant sites in the population. Curve 1 corresponds to a specific definition of what site is considered deviant. A more stringent definition of deviancy would be associated with a smaller ellipse. This will result in fewer deviant sites captured by the sieve (smaller set B) (i.e., the rarer the hunted animal, the more difficult is its capture).

The union of sets B and C is the collection of all sites selected for inspection. Curve 2 corresponds to a specific criterion by which sites are selected for further examination. A less-stringent selection criterion corresponds to the large circle. This results in a larger number of sites that require close inspection and also a larger number of deviant sites captured by the sieve.

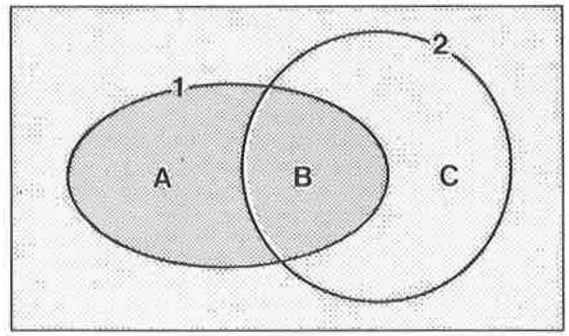

FIGURE 1 False negatives $\{A\}$, correct positives $\{B\}$, and false positives $\{C\}$.

In general, the more stringent the criterion of deviancy, the more difficult it is to identify deviant sites. The more stringent the selection criterion, the smaller the number of deviant sites captured by the sieve.

In present practice a site is considered to be a blackspot if its accident record deviates k standard deviations from the norm. The value of k is linked to statistical level of significance, and the practice in this case is borrowed from industrial quality control. What value to use for k is largely a matter of custom, with no apparent rationale. This is why it appears sensible to examine whether it is possible to discard what is arbitrary and use instead measures of performance that have clear meaning.

In this paper the focus is on blackspots that occur on a road system. This is why researchers speak of sites, road sections, ramps, intersections, and so forth. It is worth noting that identical issues arise when trying to identify deviant drivers, and that the results of analysis apply equally in both cases. To underscore this point, one example will deal with the population of drivers instead of the population of road sections.

MEASURES OF EFFICIENCY FOR A SIMPLE SIEVE

The simplest case is usually the easiest to analyze. Once the categories of thought and lines of argument for the simple case are established, the examination of more complex sieves can be undertaken.

A mathematical notation was not introduced in the first section because the central issues could be explained without burdening the reader with symbols. However, the main content of this section is analysis, and it would be inefficient to postpone the use of a precise notation any longer. Therefore, let λ be the expected (average in the long run) number of recorded accidents prevailing at a site during a specified period of time, and let x be the number of accidents actually recorded for that site and period of time.

In this section the performance of a sieve is examined, the aim of which is to identify sites for which λ is larger than some limiting value λ^*. This is done by selecting for inspection sites for which x is not less than some limiting value x^*.

For a specific site, λ is never known. What is known is x. Therefore, the question is: What can be said about the λ of a site if its x is known? The answer is best stated in terms of a conditional probability distribution. The corresponding symbol has to be added to the notational arsenal. Thus let $F(\lambda|x)$ be the probability that the expected number of accidents at a site was less than or equal to λ when the number of accidents actually recorded was x.

To provide the reader with a sense of direction, it is best to first show that $F(\lambda|x)$ is the kingpin on which everything hinges. Indeed, when $F(\lambda|x)$ is known, the performance of a sieve can be described with ease and precision. How to estimate $F(\lambda|x)$ will be described later.

The information on which analysis is based is the knowledge of x for each site. Let n(x) be the number of sites (N) that had x accidents, x = 0, 1, 2, ...

1. When sites for which $x \geq x^*$ are selected for inspection, the number of sites (S) to be inspected is

$$S(x^*) = \sum_{x^*}^{\infty} n(x) \quad (1)$$

This corresponds to the number of sites in the union of sets B and C in Figure 1.

2. When sites for which $\lambda > \lambda^*$ are considered deviant, the expected number of deviant sites (D) in the population is

$$D(\lambda^*) = \sum_{0}^{\infty} n(x) [1 - F(\lambda^*|x)] \quad (2)$$

This corresponds to the expected number of sites in the union of sets A and B in Figure 1.

3. With x^* as the selection criterion and λ^* as the criterion for deviancy, the expected number of false positives (FP) is

$$FP(x^*, \lambda^*) = \sum_{x^*}^{\infty} n(x) F(\lambda^*|x) \quad (3)$$

This is the expected number of sites in set C of Figure 1.

4. Because $S(x^*)$ corresponds to the union of B and C, whereas $FP(x^*, \lambda^*)$ corresponds to set C alone, it follows that the expected number of correct positives (CP) is

$$CP(x^*, \lambda^*) = S(x^*) - FP(x^*, \lambda^*) \quad (4)$$

This corresponds to the number of sites in set B.

5. Because $D(\lambda^*)$ corresponds to the union of A and B, whereas $CP(x^*, \lambda^*)$ corresponds to set B alone, the expected number of false negatives (FN) is

$$FN(x^*, \lambda^*) = D(\lambda^*) - S(x^*) + FP(x^*, \lambda^*) \quad (5)$$

This corresponds to the number of sites in set A.

It follows that knowledge of n(x) and $F(\lambda|x)$ will enable researchers to find all figures of merit that describe the performance of the screening process. Because n(x) is obtained from the raw data, it remains to find $F(\lambda|x)$. This is the subject of the next section.

ESTIMATION OF SIEVE EFFICIENCY

Each site of a population of sites has associated with it an unknown value λ. Regarding λ as a continuous random variable within the population of sites, let $g(\lambda)$ denote its probability density function. Furthermore, let $P(x|\lambda)$ denote the probability of recording x accidents on a site where their expected number is λ. According to Bayes' theorem,

$$f(\lambda|x) \propto P(x|\lambda)g(\lambda) \quad (6)$$

Integration of $f(\lambda|x)$ yields $F(\lambda|x)$. The coefficient of proportionality is selected to make $\int_0^{\infty} F(\lambda|x)d\lambda = 1$.

It is assumed, as is common, that accident occurrence obeys the Poisson probability law. Thus

$$P(x|\lambda) = \lambda^x e^{-\lambda}/x! \quad (7)$$

The only missing link in Equation 6 is $g(\lambda)$. The clues for the estimation of $g(\lambda)$ are hidden in the numbers n(x). Because the number of accidents recorded on a site is a reflection of λ for that site (see Equation 7), the number of sites with x accidents [n(x)] must be a reflection of the distribution of λ among all sites. This is captured by the following relationship:

$$\text{Expected proportion of sites with x accidents} = E\left\{n(x)/[\Sigma n(x)]\right\}$$

$$= \int_0^{\infty} P(x|\lambda)g(\lambda)d\lambda \quad (8)$$

The problem here is to extract the function $g(\lambda)$ from Equation 8. It is a well-researched problem [see, for example, Maritz (1)]. In consequence, it is possible to make use of results obtained by others. One specific case that appears to be of practical interest when $g(\lambda)$ is a gamma probability density function will be described in detail. This assumption is common in actuarial literature [see, for example, Buhlmann (2) or Freifelder (3)] and is used to describe the distribution of expected claim frequencies for a population of insureds. The results to follow were obtained and used by Jarrett et al. (4) when estimating the magnitude of the regression to the mean in before-and-after comparisons.

When $g(\lambda)$ is a gamma probability density function and Equation 7 holds, the probability that a site selected at random has x accidents is given by the negative binomial probability law. Therefore, the parameters of $g(\lambda)$ can be estimated easily from the sample mean and sample variance of x as follows.

1. Calculate sample mean and variance (unless indicated otherwise, summation is over all values of x):

$$\bar{x} = \Sigma\, x\, n(x)/\Sigma n(x) \quad (9)$$

$$s^2 = [\Sigma (x - \bar{x})^2 n(x)]/\Sigma n(x) \quad (10)$$

2. Estimate parameters α and β and write $g(\lambda)$:

$$\hat{\alpha} = \bar{x}/(s^2 - \bar{x}) \quad (11)$$

$$\hat{\beta} = \bar{x}^2/(s^2 - \bar{x}) \quad (12)$$

With this,

$$g(\lambda) = \alpha^{\beta} \lambda^{\beta-1} e^{-\alpha\lambda}/\Gamma(\beta), \quad \text{when } \lambda \geq 0 \quad (13)$$

By using the results of Equations 7, 8, and 13,

$$f(\lambda|x) \propto \lambda^{x+\beta-1} e^{-\lambda(1+\alpha)} \quad (14)$$

which is also a gamma probability density function with

$$E\{\lambda|x\} = (x+\beta)/(1+\alpha) \tag{15}$$

$$VAR\{\lambda|x\} = (x+\beta)/(1+\alpha)^2 \tag{16}$$

It follows that $f(\lambda^*|x)$ is a gamma probability distribution function and estimates of its parameters α and β are known. $F(\lambda^*|x)$ may be found by using numerical integration on

$$\int_0^{\lambda^*} \lambda^{x+\beta-1} e^{-\lambda(1+\alpha)} d\lambda / \int_0^{\infty} \lambda^{x+\beta-1} e^{-\lambda(1+\alpha)} d\lambda \tag{17}$$

What remains to be done is to apply these results to some actual cases.

TWO ILLUSTRATIVE EXAMPLES

The theory developed so far suffices to describe the performance of a simple screening process. The numerical examples in the following sections will serve to show what may have been obscured by convoluted mathematical arguments.

Illustrative Example 1: Ontario Highway Ramps

The second column of Table 1 lists the number of Ontario highway ramps that, in 1978, had x = 0, 1, 2,...,14 accidents. The third column lists what should be expected if, indeed, the distribution of λ is as has been assumed in Equation 13. It appears that there is satisfactory support for making this assumption.

TABLE 1 Accidents on Ontario Highway Ramps in 1978

No. of Accidents (x)	No. of Ramps with x Accidents [n(x)]	No. of Ramps Expected by Negative Binomial Model
0	2,254	2,278
1	286	249
2	95	98
3	48	48
4	21	26
5	7	15
6	8	9
7	6	5
8	5	3
9	3	2
10	0	1
11	1	1
12	0	–
13	1	–
14	1	–

By using Equations 9-12, it is shown that $\bar{x}$ = 0.3414, s^2 = 1.0677, $\hat{\alpha}$ = 0.47, and $\hat{\beta}$ = 0.16. Therefore, by using Equation 13,

$$g(\lambda) = 0.152\ \lambda^{-0.84} e^{-0.47\lambda} \tag{18}$$

This is an estimate of how λ was distributed in the population of Ontario highway ramps in 1978 and is of considerable interest by itself.

In Figure 2 the probability distribution function (PDF) of λ, based on Equation 17, is shown. It appears that 10 percent of the ramps (276 ramps) have $\lambda > 1$ accidents per year, 5 percent of the ramps have $\lambda > 1.8$ accidents per year, and so forth. It is the ramps with relatively high values of λ that may demand closer examination and that the screening procedure should identify.

The probability distribution of λ in the subpopulation of sites that had x accidents can also be

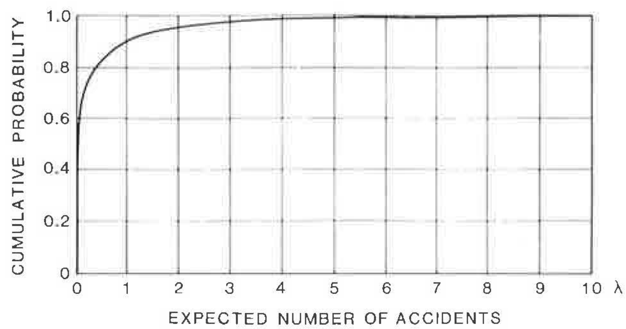

FIGURE 2 PDF of λ in the population of Ontario highway ramps (1978).

shown. This is accomplished by making use of Equations 14-16. In Figure 3, $F(\lambda = 1|x)$ for x = 1, 2, 3, 5, and 10 are shown.

Suppose that on the basis of Figure 2 researchers wish to identify those ramps for which $\lambda > 1$. There are some 276 such ramps. By using the terminology established earlier, $\lambda^* = 1$. This is shown by the vertical line in Figure 3.

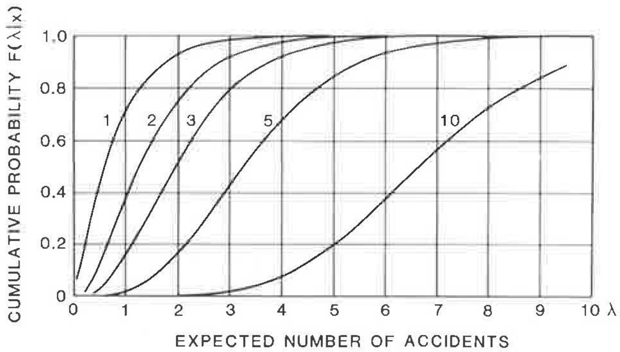

FIGURE 3 PDF of λ on ramps with 1, 2, 3, 5, and 10 accidents in 1978.

Consider now the 95 ramps that had two accidents (Table 1). From Figure 3, some 39 percent of those are expected to have $\lambda < 1$. Thus in this group of ramps it should be expected that 95 x 0.39 = 37 false positives and 95 - 37 = 58 correct positives. Similarly, the 48 ramps with three accidents each are expected to contain 48 x 0.16 = 8 false positives and 40 correct positives. Proceeding in this fashion, the data in Table 2 can be generated.

Columns 1 and 2 of Table 2 are the raw data copied directly from Table 1. The cumulation from below of the entries in column 2 yields $S(x^*)$ in column 3. Thus if sites with three or more accidents are selected for detailed scrutiny, 101 ramps have to be inspected.

The mathematical machinery assembled in the previous section and, in particular, Equations 14-16 facilitate the calculation of $F(\lambda^* = 1|x)$ in column 4. Because the computation is tedious, a FORTRAN computer code has been written for that purpose.

The products of entries in columns 2 and 4 are estimates of the number of false positives to be expected in the group of ramps that had x reported accidents, as explained earlier. Column 6 is the cumulation from below of the entries in column 5 and are therefore the estimates of the number of false posi-

TABLE 2 Measures of Performance for Ontario Highway Ramps with $\lambda^* = 1$

1	2	3	4	5	6	7	8
Recorded Number of Accidents	Number of Ramps With x Recorded Accidents	Number of Ramps With x* or More Recorded Accidents	Proportion of Sites With $\lambda<1$	Expected Number of False Positives	Cumulative Expected Number of False Positives	Expected Number of Correct Positives (3)-(6)	Expected Number of False Negatives $CP(0,\lambda) - CP(x^*,\lambda^*)$
x	n(x)	S(x*)	F(1|x)	(2)x(4)	FP(x*,1)	CP(x*,1)	FN(x*,1)
0	2254	2736	0.980	2209	2460	276	0
1	286	482	0.718	205	251	231	45
2	95	196	0.386	37	46	150	126
3	48	101	0.158	8	9	92	184
4	21	53	0.052	1	1	52	224
5	7	32	0.017	-	-	32	244
6	8	25	0.004	-	-	25	251
7	6	17	0.001	-	-	17	259
8	5	11	-	-	-	11	265
9	3	6	-	-	-	6	270
10	0	3	-	-	-	3	273
11	1	3	-	-	-	3	273
12	0	2	-	-	-	2	274
13	1	2	-	-	-	2	274
14	1	1	-	-	-	1	275

tives in the selected ramps. Thus if $x^* = 3$, then in the 101 selected ramps it should be expected that there will be 9 ramps for which $\lambda < 1$. This makes the number of correct positives equal to 92, which is the entry in column 7.

The topmost entry in column 7 is the number of correct positives in the entire set of 276 ramps. It is, therefore, the expected number of deviant sites in the population: $D(\lambda^*) = 276$. The number of false negatives (those ramps not captured by the sieve) is calculated by subtracting from 276 the entry in column 7. Thus if $x^* = 3$ in the group of 101 ramps selected for inspection, there are 92 ramps that have $\lambda > 1$, which leaves the remaining $276 - 92 = 184$ deviant ramps undetected in the population.

For ease of visual representation, the main results from Table 2 are shown in Figure 4. Thus 84 percent of the deviant sites can be captured with $x^* = 1$. But this means that more nondeviant than deviant sites are selected for close inspection and the inspection effort is large. With $x^* = 2$, the inspection effort and the number of false positives are reduced. However, almost half of the deviant ramps remain undetected. This illustrates the main tradeoffs and also describes the power and limitations of this screening process. With a small x^*, the majority of deviant sites can be identified at the cost of having to examine a large number of them in the field. Included in the selected sites will be many that are not deviant, and their inspection may be a waste of time. With a large x^*, the number of sites to be inspected can be reduced and it can also be ensured that almost all inspected sites are deviant. In this case, however, many deviant sites will not be selected for inspection.

Thus if λ^* is given as a criterion of deviancy, the analyst can trade the cost of field inspection against the penalty of leaving a deviant site undetected.

The last issue to explore is the effect of deciding on what is to be considered deviant by the choice λ^*.

There are, on average, 0.34 accident per ramp. Setting $\lambda^* = 1$, as in Table 2, defines as deviant ramps for which the expected number of accidents is about 3 times the population average. Had $\lambda^* = 1.5$ been chosen (Table 3), the number of deviant ramps is, of course, much smaller (176). Because the object of the search is now rarer, it is more difficult to capture. Thus, although for $\lambda^* = 1$, $x = 3$, $[100(48 - 8)/48] = 84$ percent of the 48 sites with $x = 3$ were correct positives, for $\lambda^* = 1.5$ the

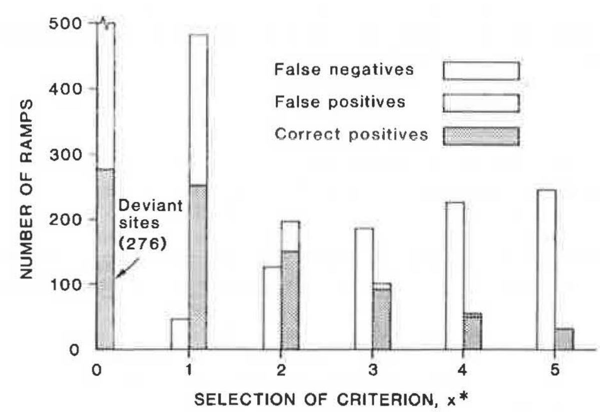

FIGURE 4 Measures of performance for $\lambda^* = 1$.

TABLE 3 Measures of Performance for Ontario Highway Ramps with $\lambda^* = 1.5$

Recorded Number of Accidents	Number of Ramps With x Recorded Accidents	Number of Ramps With x* or More Recorded Accidents	Proportion of Ramps with λ<1	Expected Number of False Positives	Cumulative Expected Number of False Positives	Expected Number of Correct Positives (3)-(6)	Expected Number of False Negatives CP(0,λ*) - CP(x*,λ*)
x	n(x)	S(x*)	F(1.5|x)	(2)x(4)	FP(x*,1.5)	CP(x*,1.5)	FN(x*,1.5)
0	2254	2736	0.993	2238	2560	176	0
1	286	482	0.859	246	322	160	16
2	95	196	0.604	57	76	120	56
3	48	101	0.343	16	19	82	94
4	21	53	0.160	3	3	50	126
5	7	32	0.070	-	-	32	144
6	8	25	0.024	-	-	25	151
7	6	17	0.007	-	-	17	159
8	5	11	0.002	-	-	11	165
9	3	6	0.001	-	-	6	170
10	0	3	-	-	-	3	173
11	1	3	-	-	-	3	173
12	0	2	-	-	-	2	174
13	1	2	-	-	-	2	174
14	1	1	-	-	-	1	175

same 84 percent yield is reached only for a larger $x = 4$.

The variation in the measures of performance of this screening process in dependence on λ^* is shown in Figure 5.

Illustrative Example 2: California Drivers

The records of 86,726 California drivers have been examined, and the number of reported accidents during 1961 have been noted (5). The number of drivers with 0, 1, 2, or 3 accidents is given in column 2 of Table 4. Column 3 gives the number of drivers in each category if the distribution of λ is as in Equation 13. This assumption is well supported. From this,

$$\bar{x} = 0.08839,\ s = 0.0939,\ \hat{\alpha} = 16.092,\ \hat{\beta} = 1.422,\ \text{and } g(\lambda) = 58.7\,\lambda^{0.4224}\,e^{-16.1\lambda}$$

On this basis, the data in Tables 5 and 6 are constructed, as in the previous numerical example. In Table 5, $\lambda^* = 0.25$, which is about 3 times the average number of accidents per driver. The difficulties of identifying deviant drivers are obvious. With $x^* = 3$, only half of those identified are deviant, yet the overwhelming majority of the 3,425

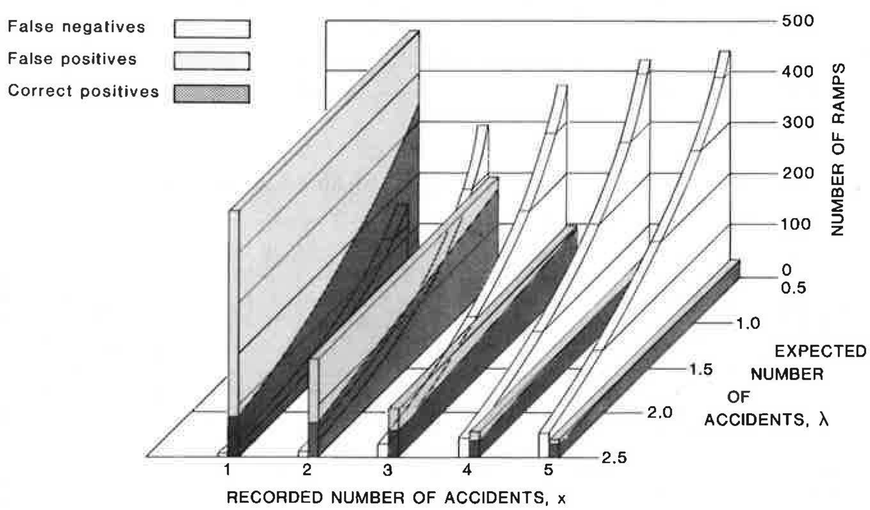

FIGURE 5 Measures of performance as a function of λ^*.

TABLE 4 Accidents to California Drivers in 1961

No. of Accidents (x)	No. of Drivers [n(x)]	No. of Drivers Predicted by Negative Binomial Model
0	79,595	79,598
1	6,638	6,624
2	451	469
3	42	31

drivers remain unidentified. It does not help much to select drivers with $x^* = 1$ because a large majority are false positives (6,205). The performance of the sieve is even worse when a more severe criterion of deviancy is considered ($\lambda^* = 0.50$ in Table 6).

SUMMARY, DISCUSSION, AND FUTURE RESEARCH

Normally, a two-stage process is used for the identification of blackspots. In the first stage a limited number of apparently dangerous locations are selected from all sites on the basis of their accident history. The sites so selected are examined in more detail in the second stage.

The data in this paper deal with the first stage of the blackspot identification process, which is likened to a sieve. A good sieve retains most sites that require detailed examination and allows through most sites that need not be examined any further.

Accordingly, a concept of sieve efficiency is proposed in which the number of sites to be inspected and the expected numbers of correct positives, false positives, and false negatives serve as measures of performance.

This concept is converted into a procedure for a special but common case, and it is applied to two illustrative examples. One deals with the population of Ontario highway ramps, and the other deals with California drivers.

In both cases the objective of the screening process is to identify units for which the expected number of accidents exceeds a given norm. What can and cannot be achieved is illustrated. Because the measures of performance are explicit, rationality in decision making and design are facilitated.

The screening process used in practice is more complex than what has been analyzed. In particular, the accident rate (accidents per vehicle kilometer), which is the most important selection criterion, is not used here. Thus the theory and computational process need to be extended so as to be applicable to the realistic blackspot identification procedures. This extension appears to be straightforward. The corresponding research work is under way.

The procedure relies on the assumption that λ obeys the gamma distribution. This may not be a good assumption in some cases. Accordingly, it is necessary to develop numerical methods to free the procedure from reliance on this assumption.

Therefore, the quality-control approach to blackspot identification does not give the analyst clues about how well or how poorly his sieve is working. In contrast, the approach suggested in this paper provides measures of performance that describe the efficiency of the sieve in intuitively clear terms.

ACKNOWLEDGMENT

The support of research by Transport Canada is gratefully appreciated.

REFERENCES

1. J.S. Maritz. Empirical Bayes Methods. Methuen and Company, Ltd., London, England, 1970.

TABLE 5 Measures of Performance for California Drivers with $\lambda^* = 0.25$ Accidents per Year

Recorded Number of Accidents	Number of Drivers With x Recorded Accidents	Number of Drivers With x* or More Recorded Accidents	Proportion Of Drivers with λ<1	Expected Number of False Positives	Cumulative Expected Number of False Positives	Expected Number of Correct Positives (3)-(6)	Expected Number of False Negatives CP(0,λ) - CP(x*,λ*)
x	n(x)	S(x*)	F(0.25 x)	(2) x (4)	FP(x*,0.25)	CP(x*,0.25)	FN(x*,0.25)
0	79595	86726	0.969	77096	83301	3425	0
1	6638	7131	0.882	5854	6205	926	2499
2	451	493	0.729	329	351	142	3284
3	42	42	0.537	23	23	19	3406

TABLE 6 Measures of Performance for California Drivers with $\lambda^* = 0.50$ Accidents per Year

Recorded Number of Accidents	Number of Drivers With x Recorded Accidents	Number of Drivers With x* or More Recorded Accidents	Proportions of Drivers with λ<1	Expected Number of False Positives	Cumulative Expected Number of False Positives	Expected Number of Correct Positives (3)-(6)	Expected Number of False Negatives CP(0,λ) - CP(x*,λ*)
x	n(x)	S(x*)	F(0.5|x)	(2) x (4)	FP(x*, 0.5)	CP(x*,0.5)	FN(x*,0.5)
0	79595	86726	0.999	79550	86648	79	0
1	6638	7131	0.996	6613	7097	34	45
2	451	493	0.985	444	484	9	70
3	42	42	0.956	40	40	2	77

2. H. Bühlmann. Mathematical Methods in Risk Theory. Springer Verlag, Berlin, 1970.
3. L.R. Freifelder. A Decision Theoretic Approach to Insurance Ratemaking. Irwin, Homewood, Ill., 1976.
4. D.F. Jarrett, C. Abbess, and C.C. Wright. Bayesian Methods Applied to Road Accident Blackspot Studies: Some Recent Progress. In Proc., Seminar on Short-Term and Areawide Evaluation of Safety Measures, Institute for Road Safety Research, SWOV, Amsterdam, Netherlands, 1982, pp. 69-74.
5. R.C. Peck, R.S. McBride, and R.S. Coppin. The Distribution and Prediction of Driver Accident Frequencies. Accident Analysis and Prevention, Vol. 2, No. 4, 1971, pp. 243-299.

Comparison of Two Methods for Debiasing Before-and-After Accident Studies

BHAGWANT N. PERSAUD and EZRA HAUER

ABSTRACT

When corrective treatment is applied to road sections, intersections, drivers, or vehicles that had a poor accident record in the past, the safety effect of the treatment is properly estimated by comparing the number of accidents in a post-treatment period with the number of accidents that would have occurred in this period without the treatment. Earlier papers have shown that simple before-and-after comparisons are consistently biased; that is, treatments appear to be more effective than they really are. Accordingly, two methods--a nonparametric method and a Bayesian method--have been separately proposed for purging this bias. The nature of the bias and the two debiasing methods are reviewed. In the main body of the paper several data sets are used to compare the performance of the methods. In most cases the Bayesian method was found to yield better estimates.

Before-and-after accident comparisons are a common method for assessing the safety effect of a treatment applied to road sections, intersections, drivers, and so forth. Conclusive evidence exists to show that when treatment is administered to systems with a poor safety record, simple before-and-after comparisons are biased (1). The bias is caused by the erroneous assumption that the number of accidents on a system in the period before treatment is an unbiased estimate of what should be expected to occur on the system during an equivalent after period had treatment not been applied. Systems with above-average accident numbers or rates in one period must be expected to show a decrease in a subsequent period even without treatment, and vice versa. This phenomenon, identified as regression-to-the mean, was demonstrated to be significant and can, in simple before-and-after comparisons, make safety treatments appear to be more effective than they really are.

To illustrate, Table 1, taken from Hauer (2), presents accident data for 20,762 1-km road sections in Ontario. Sections were grouped according to the number of accidents in 1 year. As shown by the data in the table, 12,859 sections had no accidents in that year; 4,457 had one accident, and so forth. Column 3 shows that, for each group, the average number of accidents recorded in the subsequent year revealed a reduction in the number of accidents in the second year for each group of sections with accidents in the first year. These reductions are balanced by the 12,859 sections that had no accidents in the first year but experienced an increase to 0.404 accident per section in the second year.

This is the essence of the regression-to-the-mean. When a random down-fluctuation occurs, as for the group with no accidents, an upward return to the mean for that group should be expected; when a random up-fluctuation occurs as it does for all the other groups, a downward return to the group mean should be expected.

Although there has been an increasing awareness of the phenomenon, its effect has often been dismissed because it will rarely be statistically sig-

TABLE 1 Regression-to-the-Mean: Ontario Data *(2)*

No. of Sections in Group	No. of Accidents for Avg Section in Group: First Year	Second Year	Change (%)
12,859	0	0.404	—[a]
4,457	1	0.832	-16.8
1,884	2	1.301	-35.0
791	3	1.841	-38.6
374	4	2.361	-41.0
160	5	3.206	-35.9
95	6	3.695	-38.4
62	7	4.968	-29.0
33	8	4.818	-39.8
14	9	6.930	-23.0
33	≥ 10[b]	10.390	-22.0

[a]Increase. [b]Average = 13.33.

nificant, and it is not often likely to lead to serious results (according to 1965 data from the Road Research Laboratory). Column 4 in Table 1 converts these regressions to percent changes, which show that, contrary to this opinion, the phenomenon is consistent, real, and nothing short of dramatic. Hauer (2) showed that not only road sections are subject to this phenomenon. In fact, any element of the transport system for which events occur randomly will be subject to regression-to-the-mean.

In summary, the number of accidents on a system in the before period does not, on average, remain the same in an equivalent after period. When safety treatment is applied, an estimate of the number of accidents that would have occurred in a subsequent period without the treatment needs to be made. In the next section, procedures for doing so are reviewed.

REVIEW OF METHODS FOR DEBIASING BEFORE-AND-AFTER COMPARISONS

Establishment of control groups, where possible, is perhaps the best method for obtaining estimates of what the number of after period accidents would have been without treatment. When doing so is not practical, two analytic methods are available. Details of these methods are given elsewhere (1-4).

Method 1

The nonparametric (NP) method (1,2) is simple to apply yet is based on intricate statistical reasoning. To estimate the number of accidents α_k expected to occur during an equivalent after period on a system that had k accidents in the before period, the following factors need to be known: N_k = number of systems with k accidents in the population of similar systems, and N_{k+1} = number of systems with (k + 1) accidents in the population of similar systems. Then,

$$\alpha_k = [(k+1)\,N_{k+1}]/N_k \tag{1}$$

The simple formula relies on the sole assumption that accidents on any system are Poisson distributed. Unlike the alternative method, no assumptions are made about the underlying distribution of accidents in the population of systems. To illustrate the use of Equation 1, α_3 was estimated for the Ontario data in Table 1. Here $N_3 = 791$ and $N_4 = 374$ from column 2. Thus, based on first-year data, the estimate of the number of accidents in the second year on a section that in the first year had three accidents is given by

$$\alpha_3 = (4 \times 374)/791 = 1.891.$$

This compares to 1.841 actually observed in that year.

If there is interest in estimating total accidents for cumulative groups with k or more accidents, it is not necessary to apply the nonparametric method individually for each accident group. It can be shown [see Hauer (2) for proof] that

$$A_k = N_{k+1}^{(+)} \tag{2}$$

where A_k is the estimated total number of accidents on systems that in the before period had k or more accidents, and $N_{k+1}^{(+)}$ is the total number of accidents on those systems that in the before period had (k + 1) or more accidents.

Method 2

The empirical Bayesian (EB) method (3,4) is just as simple to apply as method 1, but it is based on stronger assumptions and requires accident data for the entire population of systems. As before, it is assumed that the number of accidents for a system obeys the Poisson with a mean characteristic of the system. Furthermore, it is assumed that the distribution of these means in a population of systems can be approximated by a gamma distribution. With these two assumptions, the number of systems of a population with k accidents must obey the negative binomial distribution except for a rare situation discussed later in this section.

The expected number of accidents α_k' in the after period on a system that had k accidents in the before period is given by

$$\alpha_k' = [(k+1)\,N_{k+1}']/N_k' \tag{3}$$

N_k' is the number of systems expected by the negative binomial distribution to have k accidents. (Note the similarity between Equation 3 and Equation 1 for the nonparametric method, and recall that in Equation 1, N_k was the actual number of systems with k accidents.)

To employ this method, the before period accident data are used to get the sample mean (m) number of accidents and sample variance (s^2) for the population of systems. From these, estimates of the parameters b, c of the gamma distribution can be obtained, as follows:

$$b = m^2/(s^2 - m) \qquad m < s^2 \tag{4}$$

$$c = m/(s^2 - m) \qquad m < s^2 \tag{5}$$

As shown by Jarrett et al. (3) and by Abbess et al. (4) Equation 3 then reduces to

$$\alpha_k' = (b+k)/(c+1) \qquad m < s^2 \tag{6}$$

It should be noted that if the negative binomial distribution were to fit all of observed frequencies perfectly, then the two methods would give identical estimates.

For the rare situations when the sample mean is not less than the sample variance ($m \geq s^2$), Equations 4, 5, and 6 do not apply. Instead, the distribution of means in the population of systems approximates the limiting form of the gamma distribution, where each system has the same expected number of accidents. Therefore, instead of Equation 6,

$$\alpha_k' = m \qquad \text{for } m \geq s^2 \tag{7}$$

To illustrate the more common case, suppose again that k = 3 for the Ontario data. The sample mean of the number of accidents in the first year is m = 0.707 and the sample variance is $s^2 = 1.6491$. From Equations 4 and 5 the following estimates are obtained: $\hat{b} = 0.5345$, and $\hat{c} = 0.7540$. Therefore,

$$\alpha_3' = 3.5345/1.7844 = 2.015.$$

This also compares favorably with the observed 1.841 (Table 1). A systematic comparison of the performance of both methods is the subject matter of the next section.

In using the EB method to estimate total accidents A_k' for cumulative groups, the equivalent expression to Equation 2 is

$$A_k' = N_{k+1}'^{(+)} \tag{8}$$

where $N'^{(+)}_{k+1}$ is the total number of systems expected by the negative binomial distribution to have (k + 1) or more accidents. Recall that to get α_k in Equation 6, it was not necessary to get the N'_k's, so it may not always be convenient to apply this shortcut with the Bayesian method.

COMPARISON OF THE TWO METHODS

Given the differences between the two methods, it is of interest to compare estimates obtained by each method against what was actually recorded to see if there are circumstances in which one or the other should be preferred.

Data

Eleven primary data sets were used in this comparison. Some of the data sets contain several years of accident history, so it was possible to effectively increase the number of comparisons by varying the before and after periods. In addition, one of the driver accident data sets was disaggregated into five age groups. Thus the comparisons were done for a total of 42 data sets that involved a variety of systems (driver accidents, driver violations, road sections, intersections, and roundabouts), and that covered a variety of countries (the United States, Canada, Sweden, Israel, and the United Kingdom) and a variety of before period lengths. A total of 293 comparisons were obtained. These data sets are identified in Table 2.

Analysis and Results

To illustrate the nature of the performance comparisons, the data in Table 3 present the results for the Ontario data set. Columns 1, 2, and 6 merely repeat the data in Table 1. As shown in the first line, by using the negative binomial distribution, it would be estimated that 13,222 sections (column 3) are expected to have 0 accidents (column 1) compared with the 12,859 sections (column 2) actually counted. The nonparametric method estimates that one such section chosen at random would average 0.35 accident (column 4) during the second year, whereas by the Bayesian method the estimate is 0.31 accident (column 5). These estimates are compared to the 0.40 accident per section actually recorded in the second year. In Figure 1, 0.35 on the ordinate plotted against 0.40 on the abscissa is point A, which is designated by an empty circle; 0.31 plotted against 0.40 is point B shown by a full circle. Thus data in Table 3 yield 10 pairs of circles.

Similar tables for all of the data sets produced the data for Figures 1 and 2, where estimates from each of the two methods are plotted against what was recorded. For clarity and for reasons discussed later, the driver and the road data sets are plotted separately. In these figures the empty circles represent the nonparametric estimates, whereas the full circles plot the Bayesian estimates. Some observations follow.

For both driver and road systems, the full circles tend to hug the diagonal somewhat closer than the empty circles. Thus it is concluded that the Bayesian method is likely to give somewhat better estimates.

For the drivers (Figure 2), the nonparametric method consistently overestimates the number of accidents (or violations) per driver from about 0.2 accident per year on. In an earlier paper (1), it was speculated that this is a reflection of maturation and possibly the effect of accidents or convic-

TABLE 2 Comparison of Parametric and Nonparametric Bayesian Estimates

N - indicates the non-parametric method is better; P - indicates the parametric method is better

	χ^2		Mean		Variance		No. of Accidents										
DATASET DESCRIPTION	Before	After	Before	After	Before	After	0	1	2	3	4	5	6	7	8	9	10
N. Carolina Driver Accidents (Years 1,2 Before)	156	304	.122	.130	.143	.151	N	N	P	P	P	P	P				
N. Carolina Driver Accidents (Years 3,4 Before)	304	156	.130	.122	.151	.143	NP	N	P	P	P	P	P				
Ontario Road Sections	100	-	.707	-	1.649	-	N	N	N	N	NP	P	P	P	P	N	P
Sweden Road Junctions "Reported"	38	-	.833	-	1.946	-	N	P	N	N	P	P	P				
Sweden Road Junctions "Personal Injury"	2	-	.197	-	.393	-	P	P	P	N							
Driver Violations-North Carolina (Years 1,2 Before) (3,4 After)	550	664	.225	.252	.367	.401	N	N	N	P	P	P	P				
Driver Violations-North Carolina (Years 3,4 Before) (1,2 After)	664	550	.252	.225	.401	.367	N	N	N	P	P	P	P	P	P		
U.K. Roundabouts	5	-	3.911	-	17.245	-	N	P	P	P	P	P	P	P	P		
New Mexico Run Off Road (80,81 Before)(82 After)	55	67	.773	.383	1.747	.659	N	N	N	P	P	P	P	P	N	P	P
New Mexico Run Off Road (82 Before) (80,81, After)	67	55	.383	.773	.659	1.747	N	N	N	P	P	P	P	P	P		
New Mexico Fixed Object (80,81 Before)(82 After)	28	9	.263	.129	.459	.184	N	N	P	P	P	P	P	P	P		
New Mexico Fixed Object (82 Before)(80,81 After)	9	28	.129	.263	.184	.459	N	N	N	N	N	P	P				
North Carolina Driver Accidents (Yrs.1,2,3 Before)(Yr. 4 After)	473	74	.187	.064	.231	.070	N	N	P	P	P	P	N				
North Carolina Driver Accidents (Year 4 Before)(Year 1,2,3 Aft.)	74	473	.064	.187	.070	.231	NP	N	P	P	P	P					

TABLE 2 (continued)

N - indicates the non-parametric method is better; P - indicates the parametric method is better

DATASET DESCRIPTION	χ^2 Before	χ^2 After	Mean Before	Mean After	Variance Before	Variance After	No. of Accidents 0	1	2	3	4	5	6	7	8	9	10
North Carolina Driver Accidents (Year 1 Before)(Yrs.2,3,4 After)	21	1038	.061	.191	.066	.235	NP	N	P	P	P	P					
North Carolina Driver Accidents (Yrs.2,3,4 Before)(Yr.1 After)	1038	21	.191	.061	.235	.066	P	N	NP	P	P	P	P				
North Carolina, 22-25 year olds (Yrs. 1,2 Before)(Yrs.3,4 After)	3	26	.160	.170	.191	.203	NP	NP	N	N	P	P	P				
North Carolina, 22-25 year olds (Yrs.3,4 Before)(Yrs. 1,2 After)	26	3	.170	.160	.203	.191	N	N	N	P	P	P	P				
North Carolina, 26-39 year olds (Yrs.1,2 Before)(Yrs.3,4 After)	160	89	.127	.134	.151	.158	N	N	P	P	P	P	P				
North Carolina, 26-39 year olds (Yrs.3,4 Before)(Yrs.1,2 After)	89	160	.134	.127	.158	.151	P	N	P	P	P	P	N	P			
North Carolina, 40-59 year olds (Yrs.1,2 Before)(Yrs.3,4 After)	68	91	.107	.114	.122	.129	N	N	P	N	N	P	P	P	P		
North Carolina 40-59 year olds (Yrs.3,4 Before)(Yrs.1,2 After)	91	68	.114	.107	.129	.122	P	N	P	P	P	P	P	P			
North Carolina 60+years (Yrs.1,2 Before)(Yrs.3,4 After)	17	115	.111	.114	.125	.131	N	N	P	N	P	P	P	P			
North Carolina 60+years (Yrs.3,4 Before)(Yrs.1,2 After)	115	17	.114	.111	.131	.125	N	N	P	P	P	P	P	P			
North Carolina 21 year olds (year 1 Before)(Yrs.2,4 After)	1	1	.106	.098	.118	.108	NP	N	P	P	N						
North Carolina 21 year olds (Year 2 Before)(Year 1 After)	1	1	.098	.106	.108	.118	NP	P	N	P	N						
Israeli Road Sections (Yrs.2,3,4 Before)(Yrs.5,6,7Aft)	3	16	1.685	1.909	4.379	4.095	N	P	P	P	P	P	P				
Israeli Road Sections (Yrs.5,6,7 Bef.)(Yrs.2,3,4 Aft.)	17	3	1.909	1.685	4.095	4.379	N	P	N	P	P	P	P				

N - indicates the non-parametric method is better; P - indicates the parametric method is better

DATASET DESCRIPTION	χ^2 "Before"	χ^2 "After"	Mean Before	Mean After	Variance 'Before'	Variance 'After'	No. of Accidents 0	1	2	3	4	5	6	7	8	9	10
Israeli Road Sections (Year 1 Before)(Year 2 After)	1	3	.577	.589	.972	.943	P	P	N	P							
Israeli Road Sections (Year 2 Before)(Year 1 After)	3	1	.589	.577	.943	.972	P	P	P	N							
Israeli Road Sections (Year 6 Before)(Year 7 After)	1	2	.661	.705	1.01	1.03	N	N	P	N							
Israeli Road Sections (Year 7 Before)(Year 6 After)	2	1	.705	.661	1.03	1.01	N	P	N	P							
Israeli Road Sections (Yrs.1,2,3 Before)(Yrs.4,5 After)	7	9	1.710	1.799	3.675	3.88	N	P	P	N	P	N	P	P			
Israeli Road Sections (Yrs.4,5,6 Before)(Yrs.1,2,3Aft)	9	7	1.799	1.710	3.88	3.675	N	P	N	P	N	P	N				
Israeli Road Sections (yrs.1,2 Bef.)(Yrs.3,4 After)	4	2	1.124	1.169	2.245	2.30	P	P	P	N	P	P					
Israeli Road Sections (Yrs.3,4 Bef.)(Yrs.1,2 After)	2	4	1.169	1.124	2.30	2.245	N	P	N	N	P	N					
Israeli Road Sections (Yrs.3,4 Before)(Yrs.5,6 Aft.)	2	8	1.127	1.220	2.295	2.233	N	P	P	N	P	P					
Israeli Road Sections (Yrs.5,6 Bef.)(Yrs.3,4 After)	8	2	1.220	1.127	2.233	2.295	N	N	P	P	P	P					
Westminister Blacksites	60	-	3.223	-	19.29	-	not reported	N	P	P	P	P	P	P	P		
California Driver Accidents (72,73 Before)(74 After)	9	2	.133	.048	.149	.051	P	N	P	P	P	P					
California Driver Accidents (74 Before)(72,73 After)	2	9	.048	.133	.051	.149	NP	P	N	P							
Philadelphia Intersections (68 Before)(69 After)	3	4	.759	.797	.970	.893	N	P	N	P	P						
Philadelphia Intersections (69 Before)(68 Before)	4	3	.797	.759	.893	.970	P	P	P	P	P						

Note: N indicates that the nonparametric method is better, and P indicates that the parametric method is better. The North Carolina driver data sets are from Stewart and Campbell *(5)*, the United Kingdom roundabout data sets are from Helliar-Symons *(6)*, and the Sweden intersection data sets are from Brude and Larsson *(7)*.

TABLE 3 Application of Estimating Methods to Ontario Road Sections Data

No. of Accidents (k)	No. of Sections with k Accidents	No. of Sections Estimated by Negative Binomial Distribution to Have k Accidents	Nonparametric Estimate	Bayesian Estimate	Recorded After
0	12,859	13,222	0.35	0.31	0.40
1	4,457	4,029	0.84	0.87	0.83
2	1,884	1,762	1.26	1.44	1.30
3	791	850	1.89	1.99	1.84
4	374	428	2.14	2.59	2.36
5	160	221	3.56	3.16	3.20
6	95	116	4.57	3.73	3.69
7	62	62	4.26	4.30	4.96
8	33	33	3.82	4.87	4.81
9	14	8	5.71	5.44	6.93

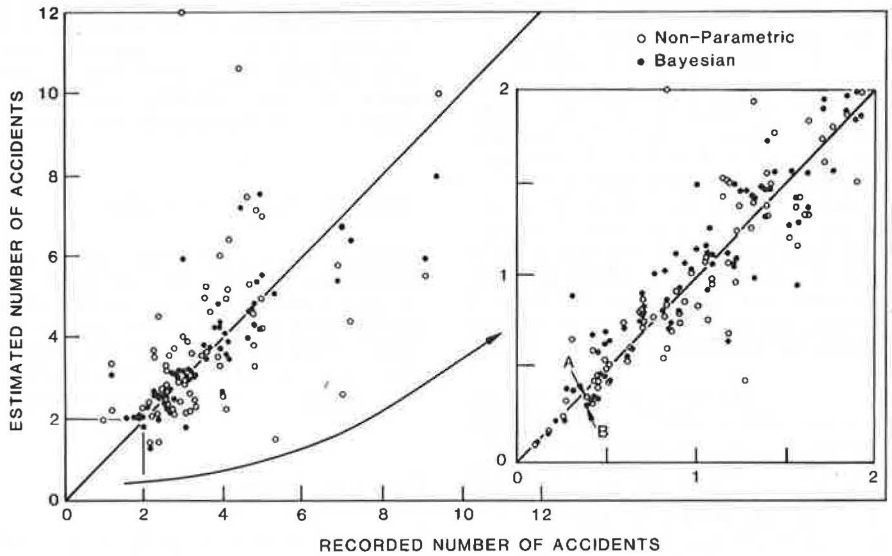

FIGURE 1 Application of debiasing methods to road data sets.

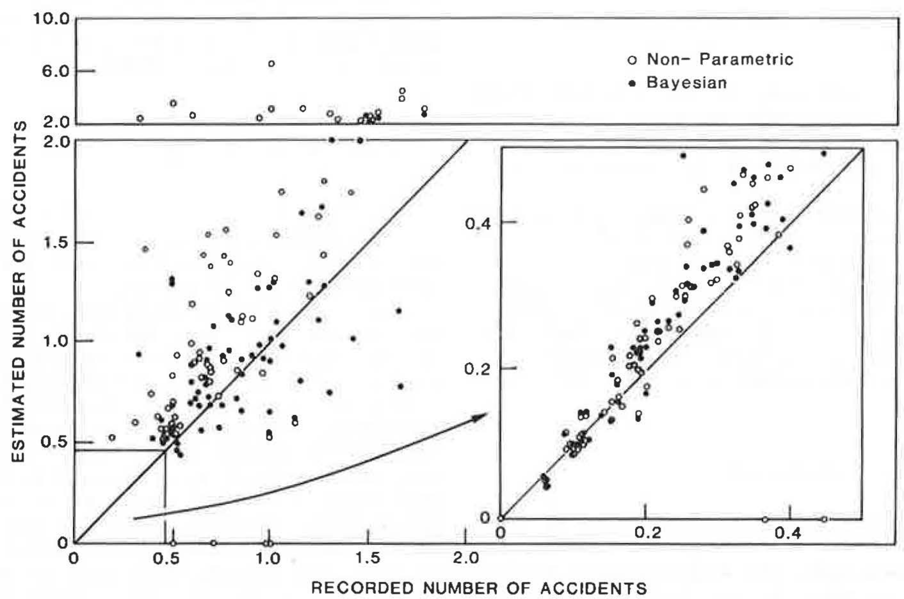

FIGURE 2 Application of debiasing methods to driver data sets.

tions on the subsequent driving record. This explanation now appears to be incorrect, for, when the order of comparison is reversed (call a later year before and an earlier year after), the same results are obtained. As indicated by the squares in Figure 2, for the Bayesian method this problem is not as severe. The overestimates are not as large or as consistent.

The data in Table 2 present the results in a different form along with some additional information to assist in the discussion that follows. To illustrate, the first line in Table 2 gives information for the entire North Carolina driver population, with the first 2 years of data representing the before period and the second 2 years representing the after period. In fitting a negative binomial distribution to the before period frequencies, a chi-square value of 156 was calculated (column 2), whereas for the after period a value of 304 (column 3) was obtained. The average driver had 0.122 accidents (column 4) in the before period and 0.130 accidents (column 5) in the after period. The sample variances associated with these two means were 0.143 and 0.151 (columns 6 and 7), respectively. For this data set, for k = 0 or 1 accident (column 8), the nonparametric estimate was closer to what was recorded, whereas for other values of k the Bayesian method was closer. For the second entry, the second 2 years of data were used for the before period with the first 2 years as the after period, and so on. Note that when the lengths of the before and after periods differ, so do the orders of magnitude of the mean numbers of accidents (e.g., the New Mexico data).

The data in Table 2 confirm that, on the whole, the Bayesian method gives better results and, in addition, reveals something that is not immediately apparent in Figures 1 and 2. For systems with k = 0 or 1, the nonparametric method performs, in most cases, at least as well as the Bayesian method. This finding has important implications, as the effect of treatments on systems with k = 0 or 1 accident is often of interest.

Discussion of Results

Effect of Type of System

From the data sets examined, it is apparent that there is a need to distinguish between road systems and driver systems. Why this is so remains an interesting research question that is currently being investigated.

Disaggregation of the data sets does not appear to have any influence on the performance of the methods. In a real application, a treatment program may be aimed at a fairly narrow group (e.g., young drivers, signalized intersections, head-on collisions). Consequently, it is important that the population be defined to include only similar systems. This issue of population definition is being researched further.

Effect of Number of Accidents (k)

For any value of k larger than 1, the number of road (nondriver) systems tends to be relatively small, so it is not surprising that the nonparametric method does not perform as well as the Bayesian method. This finding lends empirical confirmation to statements made by other researchers (1-4) about the influence of random variations in observed frequencies when the number of systems is small. By smoothing these frequencies, the Bayesian method provides more reliable estimates for the smaller groups of systems. For a more general discussion of this issue, see Maritz (8).

Effect of Number of Systems in a Group

For road systems at least, it is expected that the size of a group with k accidents would be a more direct index of the relative performance of the methods than the value of k. However, from the examination of the data, it appears that statements about the relative performance of the two methods based on group size are not clear-cut. The best that can be said about the methods is that the nonparametric method is at least as good as the Bayesian method when the number of road systems with k accidents is larger than 200. If this was made into a rule, however, there would be many exceptions. For drivers, although it appears reasonable that the size of the group must play a role in the performance of the methods, the overestimation problem prevents this issue from being examined.

Effect of Chi-Square Values

Analysis of a wide range of data sets with diverse chi-square values (see Table 2) suggests that, contrary to intuition, chi-square values for the before period data do not appear to be a good index of the performance of either of the methods. Even when the after period data also have small chi-square values, a reliable estimate is not guaranteed.

Effect of Parameters

For the Bayesian method, the sample parameters appear to be more relevant than chi-square values in determining the performance with respect to road systems. Once the sample means and sample variances for the before period data are close to these values for the after period, the Bayesian method tends to give more reliable estimates for road systems. The same conclusion cannot be made for the nonparametric method or for driver systems.

SUMMARY AND CONCLUSIONS

In this paper the regression-to-the-mean phenomenon was reviewed along with two analytic methods for purging the resulting bias from the results of before-and-after comparisons.

The focus of the paper was on an empirical comparison of the two methods: the nonparametric method where only observed accident frequencies are used to estimate the expected number of future accidents, and the Bayesian method where an assumed underlying statistical distribution smooths these frequencies before using them in estimations. The comparison, based on a large number and variety of data sets, indicated that, in general, the Bayesian method gives somewhat better estimates and should be used in assessing the safety effect of a treatment. However, for systems with zero or one accident, the nonparametric method gives slightly better results and might be preferred if the future expected number of accidents on these systems is of interest. The nonparametric method is also preferred if, in revising estimates for previous studies, accident data are only available for high-accident locations.

Discussion

Olga J. Pendleton*

In the paper by Persaud and Hauer, the authors attempt to show that, by way of example, the Bayesian method for estimating accidents in the after period is better than the nonparametric method. Whereas the data sets to which this comparison was applied appear to support this claim, it should be noted that (a) the authors do not apply any statistical methods in making the comparison and appeal only to graphical and numerical descriptive measures to support their claim, and (b) an example is not a proof.

Addressing the first comment (a), this paper would be greatly enhanced if the authors applied relatively simple statistics in making the comparison between methods. For example, along with the plots depicting the relationships of the two methods that compare actual and estimated values, statistics such as the correlation coefficient and the mean squared error of deviation from the line representing equality could be reported for the two methods and equality could be statistically tested. It also appears that there is a region of accident frequency where the comparison of these methods may yield different results (e.g., <0.5 and >0.5 in Figure 2). The nonparametric technique might even be better at <0.2. Another statistical test that could be made is a simple t-test on the differences of the methods or for a more nonparametric approach, a χ^2 test of observed versus expected for each method. These statistics would be easy to apply and lend more credence to the authors' claims.

The second comment (b) is motivation for future research. Either a rigorous mathematical proof that compares the power of the two techniques or a simulation study would be interesting. In light of the difficulty of this task, this paper did a suitable job of attempting to answer this question in a less rigorous but informative and interesting manner.

*Operations Research and Analysis Program, Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843

Authors' Closure

We thank Pendleton for her interest in our paper and for highlighting an apparent shortcoming. We think that the shortcoming is not so much in the omission of statistical tests as in the absence of any rationalization of this omission.

After reviewing our results, we chose not to perform any statistical tests as we felt that the conclusion that "the Bayesian method is likely to give somewhat better estimates" was ably supported by the plots. We did not seek a stronger conclusion because, as Pendleton notes, an example is not proof and, equally important, because a stronger conclusion would have detracted from our findings with regard to other issues. The effects of the type of system and the number of before accidents are issues well in keeping with our original intention "to see if there are circumstances in which one or the other method should be preferred."

In summary, although we agree in principle with Pendleton's proposals, we believe that they are outside the scope of this paper.

REFERENCES

1. E. Hauer and B. Persaud. Common Bias in Before-and-After Comparisons and Its Elimination. In Transportation Research Record 905, TRB, National Research Council, Washington, D.C., 1982, pp. 164-174.
2. E. Hauer. Bias-by-Selection: Overestimation of the Effectiveness of Safety Countermeasures Caused by the Process of Selection for Treatment. Accident Analysis and Prevention, Vol. 12, No. 2, June 1980.
3. D.F. Jarrett, C. Abbess, and C.C. Wright. Bayesian Methods Applied to Road Accident Blackspot Studies: Some Recent Progress. In Proc., Seminar on Short-Term and Areawide Evaluation of Safety Measures, Institute of Safety Research, SWOV, Amsterdam, Netherlands, 1982.
4. C. Abbess, D. Jarrett, and C.C. Wright. Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the "Regression-to-the-Mean Effect." Traffic Engineering and Control, Vol. 22, No. 10, 1981, pp. 535-542.
5. J.R. Stewart and B.J. Campbell. The Statistical Association Between Past and Future Accidents and Violations. Highway Safety Research Center, University of North Carolina, Chapel Hill, Dec. 1972.
6. R.D. Helliar-Symons. Yellow Bar Experimental Carriageway Markings: Accident Study. TRRL Report 1010. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1981.
7. U. Brude and J. Larsson. The Regression-to-the-Mean Effect--Some Empirical Examples Concerning Accidents at Road Junctions. In Proc., Seminar on Short-Term and Areawide Evaluation of Safety Measures, Institute of Safety Research, SWOV, Amsterdam, Netherlands, 1982.
8. J.S. Maritz. Empirical Bayes Methods. Methuen and Company, Ltd., London, England, 1970.