C. 20

1185

TRANSPORTATION RESEARCH RECORD

Methods for Evaluating Highway Improvements

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

Transportation Research Record 1185

Contents

Foreword

This Record is a compendium of papers related by their focus on evaluating highway improvements. They range from consideration of the effect of incomplete accident reporting on estimating safety to the methodological issues of evaluating the effect of traffic signals on traffic safety.

Hauer and Hakkert assess the magnitude of how many reportable traffic accidents are in fact reported and what the effect of this incomplete reporting is on the estimation of road safety. They examine the factors that affect the probability of an accident's being reported and develop a relationship between proportion of accidents reported and the accuracy with which the proportion is known.

Bowman and Brinkman report on a research effort to determine the effectiveness of low-cost countermeasures (combinations of signs, pavement markings, and delineators) in reducing accident potential at narrow bridges. The operational-based effectiveness evaluation was performed by conducting before-and-after analyses on vehicle speed and lateral placement at 18 narrow bridge approach sites.

Higle and Witkowski investigate the application of a Bayesian analysis of accident data to identify hazardous locations. They go on to compare the identification results from the Bayesian analysis with results that were obtained from classical statistical techniques.

The next two papers deal with traffic safety evaluation of signalized intersections. Persaud reports on a review of previous studies that evaluated the impacts of a traffic signal installation on safety. He finds that because of common statistical pitfalls, erroneous conclusions related to the safety effectiveness of traffic signals have too often been made. Recommendations on how to improve analysis are given. Hauer et al. have developed models to estimate the safety of a signalized intersection on the basis of traffic flow and accident history. The authors report that with these models, if traffic flows and accident history are available, a determination can be made on whether an intersection is unusually hazardous.

Finally, Hakkert and Pistiner discuss the use of behavioral-oriented criteria to assess the environmental quality and safety of residential streets. Their methodology was used to identify street design and traffic management techniques that have the potential to modify the traffic impacts and also positively affect the perception of environmental quality of the street by the residents.

Extent and Some Implications of Incomplete Accident Reporting

E. HAUER AND A. S. HAKKERT

Three questions are addressed: How many reportable accidents are in fact reported? What is the relationship among road safety, accidents occurring, and accidents reported? What are the effects of incomplete reporting on the estimation of road safety? A review of 18 studies made at different times and localities reveals considerable variability in the degree of nonreporting. As a ballpark estimate, fatalities may be known to an accuracy of about 5 percent. Some 20 percent of injuries requiring hospitalization and perhaps 50 percent of all injuries are not found in official statistics. Furthermore, the probability of an accident's being reported depends on the age of the victim; whether the victim is the driver, the passenger, or a nonoccupant; the number of vehicles involved; and several additional factors. Analysis shows that the accuracy with which road safety can be measured depends on the proportion of accidents reported and on the accuracy with which this proportion is known. It appears that the variance of the estimate of safety (or of the safety effect of some measure) is inversely proportional to the square of the average proportion of accidents reported.

Much of what we know and do about road safety is tied to the use of accident information reported to the police. It is therefore natural to think that if the level of accident reporting to and by the police is to be diminished, the ability to manage road safety will also be harmed. Put simply, with fewer accidents reported it will take longer to accumulate the same amount of data; therefore, black spots will take longer to detect, patterns of accidents will be more difficult to discern, the effect of interventions will be even less precisely known, and so on. It is repercussions of this kind that seem to be at the root of the concern about reduced levels of accident reporting.

To retain the proper perspective, it is important to remember that at present not all accidents are reportable and not all reportable accidents are in fact reported. Thus, even in the past we have had to make do with only a part of the accident information. Is the problem that any further decrease in accidents reported is critical? Furthermore, in many countries only injury accidents are reported to the police. Do these countries do a less creditable job of managing safety? Answers to such questions are not easy to provide.

E. Hauer, Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada. A. S. Hakkert, Transportation Research Institute, Technion City, Haifa 32000, Israel.

It appears that the discussion about the level of accident reporting might benefit from the resolution of three questions:

1. How many reportable accidents are in fact reported?
2. What is the relationship between road safety, accidents *occurring*, and accidents *reported*?
3. What are the statistical repercussions of incomplete reporting?

These three questions are dealt with in the three sections that follow. We hope that the provision of factual information (next section), conceptual clarification (third section), and analytical tools (fourth section) will bring about more enlightened debate.

KNOWLEDGE ABOUT THE EXTENT OF UNDERREPORTING

Some prefer to call it a motor vehicle accident; others call it a crash. In principle, everybody has in mind the same kind of event. In Ontario, motor vehicle accidents are defined as either a collision of a motor vehicle with a movable or fixed object or an explosion, submersion, or rollover of a motor vehicle. Motor vehicle accidents are reportable if they entail an injury (visible to the police or complained of by the victim) or if the property damage exceeds a certain limit, which is adjusted from time to time. Reportable motor vehicle accidents are events with fuzzy edges. Much of the fuzziness is due to the criteria that make an accident reportable.

In Figure 1 we show a 21-year history of police-reported accidents in Ontario. The intent is to demonstrate how rubbery the yardstick of "reportable motor vehicle accidents" is. Note that the monetary limit for reportable property-damage-only (PDO) accidents was adjusted in 1970, 1978, and again in 1985. As can be expected, some accidents that would have been reportable before the change are not reportable with the higher limit in place. The corresponding precipitous drops in the time series of PDO accidents shown in Figure 1 are evident. If this is how things work, one should also expect that as long as the reporting limit remains constant while the cost of repairs keeps rising, more and more accidents become reportable. Thus, under such conditions, the number of

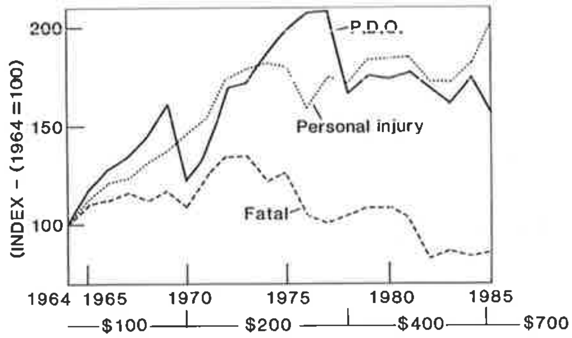

FIGURE 1 Time series of property-damage-only, injury, and fatal accidents in Ontario.

reportable accidents will increase as time passes even if there is no increase in either the number of accidents or their severity. The inflation that eats away at the value of the dollar causes an inflation of reportable accidents. Such accident inflation has nothing to do with changes in road safety. At this time it is unclear how much of the time trend in the PDO accidents in Figure 1 is attributable to "accident inflation."

The second line of defense is the injury accidents, which should be immune to the ravages of monetary inflation. The third line of defense is the fatal accidents, the count of which is generally thought to be the most reliable of all.

Unfortunately, the injury-accident yardstick is also rubbery, though for different reasons. The data in Table 1 were extracted from the Canadian national summary of motor vehicle accident casualties. The corresponding data for Ontario are shown in Table 2. The "Index" column illustrates our point.

Inspection of the "Index" column in both tables reveals that, per injury, only half as many persons die today as did 20 years ago. This trend is quite consistent, even though the number of victims killed and injured goes up and down during these two decades.

The decline in the probability of dying from an injury sustained in a motor vehicle accident (see the killed/injured ratio in Tables 1 and 2) is affected by many factors. Among the main ones that come to mind are less injurious vehicles, use of seatbelts and helmets, forgiving roadsides, and more travel in urban areas. However, the decline may be partly due to an increased inclination to report consequences of an accident as injuries. Indeed, a more detailed examination of the Ontario data revealed that proportionately more and more slight injuries are being reported to the police. Thus, although the number of injuries classified as "major" remained stable over the last decade, the number of "minimal" injuries increased by some 30 percent. Additional reasons for the observed decline in the killed/injured ratio may be that people injured in motor vehicle accidents are more resilient than they were 20 years ago, they get to the hospital in better condition because of improved emergency medical services, and more of them can be saved if they arrive in the hospital alive and are kept alive beyond the 30-day limit, which removes them from the count of motor vehicle accident fatalities.

TABLE 1 MOTOR VEHICLE ACCIDENT CASUALTIES: CANADA

Year	Victims killed	Victims Injured	Killed/Injured	Index
1965	4,902	150,612	0.0325	1.00
1970	5,080	178,501	0.0284	0.87
1975	6,061	220,941	0.0274	0.84
1980	5,461	262,977	0.0208	0.64
1984	4,120	237,455	0.0174	0.53

TABLE 2 MOTOR VEHICLE ACCIDENT CASUALTIES: ONTARIO

Year	Victims killed	Victims Injured	Killed/Injured	Index
1965	1,611	60,917	0.0264	1.00
1970	1,535	75,126	0.0204	0.77
1975	1,800	97,034	0.0186	0.70
1980	1,508	101,367	0.0149	0.56
1983	1,204	91,706	0.0131	0.50

Thus, we have cast some doubt on the reliability of the count of injury accidents reported to the police as a measure of safety. Inasmuch as the count seems to depend on the inclination to complain of injury, it may be the inclination to complain that determines the count. As an aside, we have perhaps raised a question about the count of fatalities (or fatal accidents) as a yardstick of road safety. Suppose that everything remains the same, but doctors keep more victims alive beyond the 30-day limit and therefore the count of fatalities diminishes. Do we equate advances in medicine with improved road safety? Perhaps we should.

In summary, the count of reportable motor vehicle accidents is related to road safety but cannot be considered a good measure of it. The problem of dealing with a rubbery yardstick is further compounded by the fact that not all that is reportable is also reported. A review of what is known about the extent of underreporting is the main object of this section.

Data on accidents recorded by the police are the most widely used source of information in road safety, especially for highway and traffic engineering purposes. The problem of underreporting has received a fair amount of research attention. In Table 3 the results of several studies are summarized (*1–18*). Most estimates of the proportion of accidents reported to the police come from comparisons of police data and hospital files. In some cases the comparison is among several sources (police, hospitals, fire departments, insurance companies, institutes of pathology, self-reporting, employer records, etc.).

TABLE 3 SUMMARY OF STUDIES ON UNDERREPORTING OF ACCIDENTS

Reference & Year of publication	Country	Year	Type of Study	Sample Size	% Reported
(1) 1966	U.S.A California	1963	Police vs. employee records for DOH vehicles	Small 438 cases	Fatal 100% Injury 93% PDO 38% All 49%
(2) 1969	U.S.A Mississippi		Telephone interview vs. police	500 cases	All crashes 42%
(3) 1971	U.S.A	1974	Police vs. self reported accid.		All 35%
(4) 1973	U.K.		Police vs. hospital		Serious Injury 84% Slight Injury 33% Single-vehicle Injury to driver 20%
(5) 1974	U.S.A N.C.		Insurers vs. Dept. of Motor Vehicles		All crashes 89%
(6) 1977	Canada	1974	Police vs. hospital	medium 1008 cases	Fatal 100% In-patient 97% Out-patient 76% All injuries 88%
(7) 1979	U.S.A N.D.		Insurers vs. Dept. of Motor Vehicles (motorcycles)		All crashes 47%
(8) 1981	U.S.A.	1981	Telephone Interview on unreported acc.	large 7624 hshlds.	Injury acc. 79% PDO 54%
(9) 1983	Canada	1981	Police vs. hospital Children less than 15 years old	medium 1767 cases	Injury acc. 59%
(10) 1983	U.K.	1972	Police vs. hospital	large 7630 cases	Injury acc. 50%
(11) 1984	West Germany	1980	Police vs. hospitals, fire dept., patholo-gical inst.	medium	Fatalities 91%
(12) 1984	West Germany	1980	Police vs. medical records of fatalities	medium	2-10% die after 30 days
(13) 1984	several countries	1970s 1980s	Police fatalities vs. health authority death cert.	large	Netherlands 106% New Zealand 97% Norway 80% Sweden 93% USA 96% W. Germany 104%
(14) 1984	Netherlands	1979	Police vs. hospital	large 25,000 cases	In-patients 82-85% All injuries ~45%
(15) 1985	West Germany	1978	Police vs. hospital and insurance data	medium 780 cases	all injuries ~50%
(16) 1985	U.S.A Ohio	1977	hospital vs. Dept. of Motor Vehicles	medium 882 cases	all injuries 55% drivers 74% young (<16 yrs) 28%
(17) 1985	U.S.A. California	1981/ 1982	hospital vs. police non-crash injuries (stops, swerves,..)	small 46 cases	all injuries 38%
(18) 1985	West Germany	1983	Police vs. hospitals insurance, garages	large 2744 cases	fatalities 95% serious injury 78% slight injury 62% major PDO 42%

Inspection of the entries in Table 3 leads to several observations. It seems that different studies yield widely discrepant estimates of the proportion of reportable accidents found in police records. Part of the discrepancy stems from the diversity in methods of study, another part from genuine differences associated with time and place.

In spite of the discrepant estimates, the factors that affect the inclination to report an accident emerge with clarity. Thus, it seems evident that fatal accidents are reported more fully than serious injuries and that the coverage of the latter is, in turn, better than that for slight injuries. [It deserves noting that even the count of fatalities is not without error (*12*).] As a ballpark average of the entries in Table 3, police records miss some 20 percent of injuries that require hospitalization and perhaps up to half of the injuries that do not. In addition, the probability of reporting an injury sustained in a motor vehicle accident increases with the age of the injured person. For young children it is 20 to 30 percent, and for persons over 60 it is around 70 percent (*15*). Similarly, the probability of reporting such an injury by those involved is largest for the driver, less for the passenger, and even less for nonoccupants. Another factor that affects the inclination to report an accident is the number of vehicles involved. Smith (*1*) reports 57 and 12 percent reporting levels for single-vehicle injury and PDO accidents, respectively; for multivehicle accidents the corresponding percentages are 96 and 41.

The problem of nonreporting is compounded by a variety of inaccuracies and errors that creep into the eventual computerized record of the accident. Shinar et al. (*19*) compared police records of 124 accidents with detailed information collected by multidisciplinary accident investigation (MDAI) teams. Police data were found most reliable for location, date, day of week, and number of drivers, passengers, and vehicles involved and least reliable for information about vertical alignment, road surface, and accident severity. Frequent errors were found in driver age (11.6 percent) and vehicle model year (5.3 percent). Hautzinger et al. (*15*) found inaccuracies in the reporting of accident type and vehicle maneuver in 5 to 16 percent of the cases. Kelman (*20*) finds 2.2 discrepancies between the reported and the coded information per accident. Hakkert and Hocherman (*21*) find 1 such discrepancy per accident in rural areas and 1.5 in urban areas. In particular, there are many errors in the coding of locations (25 to 39 percent). The "unknown" code appears in up to 25 percent of the cases. Keller (*22*) finds small errors (2 to 4 percent) in year of birth, sex, type of road user, number of vehicle users, road condition, and light and weather conditions. Larger errors occurred in accident location (5 to 7 percent) and still larger errors in type and cause of accident (7 to 16 percent).

The main points made in this section are as follows:

- The criteria that make a vehicle accident "reportable" make the count of reportable accidents a questionable measure of road safety. The number of collisions and their objective severity may be constant and still the number of reportable accidents may vary as a function of the limit on property damage, the inclination to complain of injury, and other extraneous factors.
- Only a part of all reportable accidents ends up in the official records. The proportion of reportable accidents in official files depends on many factors. The main determinants of the probability of finding a reportable accident reported are severity, age of victim, role (driver, occupant, etc.), and number of vehicles involved. We know little about how the proportion of accidents reported varies in time or from site to site.

CLEARING THE CONCEPTUAL UNDERBRUSH

Now that we have established that the underreporting of accidents is a sizable problem and before we proceed to explore the statistical implications thereof, we have to get our thoughts straight on matters of definition and principle.

Consider a certain road system for which the 1986 official records show 40 single-vehicle accidents in which the driver was young and impaired. In reality, say, 200 such reportable accidents occurred but only 1 in every 5 gets to be reported, on the average. Few would disagree that it is the 200 accidents that occur (not the 40 accidents that are reported) that should form the basis of statements and inferences about the safety of this road system. Although few would disagree, it is also true that almost all statements now made about road safety (especially in highway and traffic engineering) are about accidents that are reported. The distinction between accidents reported and accidents occurring is just not common in practice.

One reason for accepting this obvious paradox is the hope that as long as the level of accident reporting remains constant, it will still be possible to make comparisons. That is, it will be possible to judge improvement or deterioration and identify high-accident locations or characteristic patterns of accident occurrence even when only a portion of the accidents is reported.

Thus, to continue the above example, if 50 single-vehicle accidents by young, impaired drivers were reported on the same road system in 1985 and if there were no change in the inclination to report such accidents, one could legitimately estimate that from 1985 to 1986 the number of such accidents had decreased by $100 \times (50 - 40)/50 = 20$ percent. Note that it is not correct to say that there was a decrease of 10 accidents; for all we know, the reduction is perhaps 50 accidents (because only 1 in 5 is reported, on the average). Only statements about relative magnitude are legitimate. However, should the inclination to report an accident change from 1985 to 1986, even the possibility of making relative comparisons would be destroyed. In that event, the 20 percent reduction in the count of reported accidents could reflect a change in the level of reporting or in the number of accidents occurring, or, most likely, an inseparable mix of the two.

We are now getting close to the heart of the matter. If the inclination to report an accident is constant from time period to time period or site to site, comparisons of safety on the basis of reported accidents are legitimate. In this case, the net effect of a sudden reduction in the level of accident reporting would be that (after some transition period) it will take a proportionately longer time to collect a fixed amount of accident data. No further, complicated analysis is necessary.

However, even if present practice takes no cognizance of the distinction between accidents reported and accidents occurring, the assumption that the average ratio of "reportable accidents reported to reportable accidents occurring" is constant does not seem to be realistic. On the basis of what little we know, the probability of a reportable accident's being reported depends on a host of factors (accident severity, structure of insurance premiums, age and sobriety of driver, age of victims, inclination to seek compensation, number of vehicles involved, proximity to police station, workload of the police force, and so on). Many of these factors change with time and location.

If the inclination to report an accident cannot be realistically taken as constant, one must squarely face the fact that to make an accurate statement about road safety, one has to have some idea of the average proportion of accidents that are reported. At this point we have to state what the phrase "road safety" means.

For the purposes of science and management, "road safety" is a real characteristic; it has a magnitude and is subject to measurement. Thus we may inquire, for example, about the (magnitude of) safety of the Interstate highway system in 1980, the safety of the intersection of High and Main streets in May 1986, or the safety of John Smith next Tuesday. We will use the word "entity" to designate these elements of the real world (highways, intersections, drivers) whose safety is to be determined. The safety of an entity is defined as the number of accidents in several classes *expected* to occur on that entity per unit of time. The term "expected" means "what would be the average in the long term were it possible for all conditions to remain unchanged indefinitely." Thus, a distinction is created between accidents *expected to occur* and accidents *occurring* on an entity. It is the former that we call "safety" and it is the latter that, were it known to us, would allow us to make informed guesses (statistical inferences) about safety. However, as has been stressed, not all reportable accidents that occur are in fact reported. Thus, the link between what is occurring and what is expected to occur is severed, and the customary flow of statistical inference is obstructed.

It should be clear that without some notion about what proportion of accidents makes it into official records, one cannot establish any functional relationship between road safety and reported accidents. Thus, for example, if on a road system there are official reports of 50 single-vehicle accidents in which the driver is young and impaired and we think that 10 to 30 percent of all such accidents get reported, we can estimate that perhaps 500 − 167 such accidents have occurred. But without stating the 10 to 30 percent range, official accident records are apples and the estimates of the expected number of accidents occurring are oranges.

In summary, the link between road safety and the number of (reportable) accidents occurring is through the laws of probability and the methods of statistics. The link between (reportable) accidents occurring and their subset—accidents reported—is through the probability that a reportable accident will get into the official records. A realistic analysis of the repercussions of incomplete accident reporting must take into account the interaction of these three elements and their linkages.

In the next section we make an attempt at analysis. We face the usual dilemma. Some readers who have an interest in the results may find the machinery of analysis obscure. With this difficulty in mind, an attempt will be made, where possible, to translate mathematical statements into their real-world equivalents.

ANALYSIS

In road safety management, official accident reports are used in many ways: to keep tabs on trends, to identify target groups (accidents types, high-risk drivers, dangerous vehicles, hazardous sites, etc.) that for one reason or another demand attention, to examine the relationship between accident occurrence and various causal factors, and to examine changes in road safety as well as the reasons for such changes (as, for example, when the effect of some countermeasure is estimated or when the causal factors of changes in accident occurrence are investigated). All these diverse uses of official accident reports are linked to two generic questions about road safety:

1. What is the magnitude of road safety for some specific "entity" during a certain period of time?
2. What is the change in the relative magnitude of the safety of an entity from one period of time to another or the relative difference in the safety of several entities?

In the interest of clarity, the two cases (estimation of the magnitude of road safety and estimation of change or difference in the relative magnitude of road safety) will be treated separately.

We now introduce the requisite notation. For purposes of analysis, the safety of an entity during a specified period of time is the vector $\langle m \rangle$ of the expected number of accidents $m_1, m_2, \ldots, m_i \ldots, m_n$ in classes $1, 2, \ldots, i \ldots, n$. Thus, for example, the entity may be a specific intersection and the classes could be accident types by initial impact (rear end, head on, . . .), by accident severity (PDO, injury, or fatal, or perhaps AIS1, AIS2, . . .), or by any other category.

Answers to questions about the magnitude of the components of m (as, for example, in the identification of black spots or deviant drivers) or about changes in the magnitude of $\langle m \rangle$ (as, for example, in research about the safety effect of certain treatments) are provided with the aid of statistics. The mathematical point of departure is a functional relationship among the number of accidents reported in class i (x_i), the probability that an accident in class i will be reported (p_i), and the number of accidents expected to occur (m_i). This functional relationship will, in turn, determine how information about accidents as reported by the police $\langle x \rangle$ is used to make inferences about road safety $\langle m \rangle$. The same functional relationship among x_i, p_i, and m_i has to be used in our attempt to describe in numbers the deterioration in the knowledge of road safety that is caused by a degradation in the level of accident reporting.

If all accidents that occur were also reported, we would be on solid ground. Standard statistical literature gives guidance on how to estimate what is *expected* to occur on the basis of what has been *observed* to occur. In our case, however, only a certain portion of what occurs enters into the data. Important aspects of this problem may be found in the statistical literature under the name "partial ascertainment." It has been shown (*23–25*) that if p_i is the probability for an accident of class i to be reported and that if accident occurrence follows the Poisson (or binomial or negative binomial) probability law with m_i as the expected value, the number of *reported* accidents follows the same probability law but with ($p_i m_i$) as the expected value.

Thus, for example, if the count of accidents per year for some intersection is Poisson distributed with an expected value (mean) of 10 accidents and the probability for such accidents to be reported is 0.7, the count of reported accidents for that intersection is also Poisson distributed with an expected value of 7 accidents.

There is some good and some bad news in this message. The good news is that the loss of data does not alter the shape of the probability distribution and does not increase its variance. The bad news is that from the count of reported accidents, one can only make inferences about the expected number of reported accidents ($p_i m_i$) but not about p_i or m_i separately, a point already stressed by both Rao (*23*) and Kemp (*24*). In other words, the count of reported accidents cannot tell us anything about the number of accidents expected to occur when we have no idea what p_i is. Evident as all this is, the importance of knowing the proportion of accidents reported to the police by accident class does not seem to have been widely recognized or researched.

To make inferences about the magnitude of the components of vector $\langle m \rangle$, one has to have an estimate of p_i for all i. Barring that, it is possible to make inferences about *relative* magnitude of the m's if one is willing to assume that the same vector of values $\langle p \rangle$ applies to all entities among which the comparison is made. In other words, one can still calculate the percentage of change in accidents over two time periods (say, in a before-and-after study) or even compare the relative safety of entities in different parts of the city, state, or country if it is correct to assume that the same vector of probabilities of accidents to be reported applies to all these entities.

Estimation of Magnitude of $\langle m \rangle$

We can use data about x_i—accidents of class i reported to the police—to obtain an estimate of $p_i m_i$—the expected number of accidents in class i reported to the police. To underscore the fact that the two components of the product $p_i m_i$ cannot be told apart, we will use the notation $r_i = p_i m_i$. Thus, r_i is the expected number of *reported* accidents and $\hat{r}_i$ is its estimate. (In what follows we will use a caret to mean "the estimate of.") We are not interested in $\langle r \rangle$ per se; what we wish to estimate is $\langle m \rangle$; $\langle r \rangle$ is only a stepping-stone en route to this goal. To estimate m_i we have to make use of

$$m_i = r_i / p_i \quad (1)$$

replacing r_i by $\hat{r}_i$ and p_i by $\hat{p}_i$. Inasmuch as $\hat{r}_i$ is a function of x_i, Equation 1 is the embodiment of the functional relationship between the x's and the m's of which we spoke earlier. It is now also explicit that $\hat{p}_i$ is an essential part of this relationship; that without some knowledge about the magnitude of $\langle \hat{p} \rangle$, $\langle \hat{m} \rangle$ cannot be related to $\langle \hat{r} \rangle$ and thus no link can be established between safety and the count of accidents reported to the police $\langle x \rangle$.

Note that $\hat{p}_i$ is itself an estimate and is surrounded by uncertainty (which is at present considerable). For didactic reasons we will proceed in two steps. First we will assume that p_i is known with certainty. Next we will explore the consequences of the uncertainty surrounding the actual magnitude of p_i.

Suppose then, first, that p_i is known to us with certainty. We also know that on a particular entity, x_i accidents have been reported. What we are after is the accuracy with which m_i can be estimated from these data. The accuracy of the estimate of m_i is described by $\mathrm{Var}\{\hat{m}_i\}$. In our case,

$$\mathrm{Var}\{\hat{m}_i\} = \mathrm{Var}\{\hat{r}_i\} / p_i^2 \quad (2)$$

A numerical example might help to elucidate the meaning of Equation 2. Suppose that six injury accidents were reported at an intersection during the 3-year period 1982–1984 and it is known that 70 percent of all injury accidents are reported, on the average. In this case, $\hat{m}_{\text{injury}} = 6/0.7 = 8.6$ injury accidents for these 3 years and $\mathrm{Var}\{\hat{m}_{\text{injury}}\}$ is estimated as $6/0.49 = 12.2$ (injury accidents)^2. When the Poisson model is applied to accident occurrence, the variance of the estimate of the mean is the same as the mean. Note that when accident reporting is not complete, the variance of the estimate is always larger than the mean even if p_i is known precisely!

To elaborate on Equation 2, assume that we have n annual counts of reported accidents for some entity. If r_i is now the expected number of reported accidents per annum, $\mathrm{Var}\{\hat{r}_i\} = r_i/n$. Let m_i be the expected number of accidents per annum for this entity. Using Equation 1, it follows that $\mathrm{Var}\{\hat{r}_i\} = m_i p_i / n$. Therefore, Equation 2 takes the form $\mathrm{Var}\{\hat{m}_i\} = m_i/(np_i)$. In this form some of the trade-offs can be made visible.

In Figure 2 we show how the variance-to-mean ratio of the estimate of m changes as a function of the number of

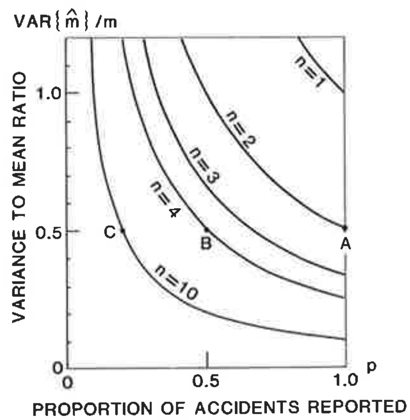

FIGURE 2 Accuracy of estimation when p is known exactly.

years of reported accident data used in the estimate and on the level of accident reporting when p is assumed to be known precisely. If accident reporting is complete ($p = 1$) and we wish the variance of the estimate of m to be half of the mean, 2 years' worth of reported accidents is needed (Point A in Figure 2). To keep the same accuracy of estimation when only 50 percent of the accidents are reported, 4 years of data are required (Point B) and with p as low as 0.2, 10 years of data would be required to retain the same accuracy.

The thrust of the argument embodied in Equations 1 and 2 and Figure 2 is as follows: the accuracy with which we can measure (estimate) road safety $\langle m \rangle$ depends not only on the count of reported accidents $\langle x \rangle$ but also on the knowledge of the proportion of accidents that is reported $\langle p \rangle$ (and, as will be shown shortly, on the accuracy with which p is known). We measure the accuracy with which we can estimate $\langle m \rangle$ by $\text{Var}\langle \hat{m} \rangle$. When for an accident class i we manage to express $\text{Var}\{\hat{m}_i\}$ as a function of p_i (and, later also, of the variance of p_i), we can examine quantitatively how the lesser reporting of accidents degrades the accuracy with which we can measure road safety. This is what we set out to do: to tell in numbers how lower levels of accident reporting affect road safety management, which is (or should be) predicated on the measurement of road safety.

Having established the pattern of analysis for the simple but unrealistic case in which p_i is known accurately, we can proceed to the more complex but perhaps more practical case in which we admit that p_i is not known to us with certainty. What we have is an estimate $\hat{p}_i$ of p_i. As with all estimates, $\hat{p}_i$ has a variance $\text{Var}\{\hat{p}_i\}$. Now both the numerator and denominator in Equation 1 are replaced by estimates that have to be regarded as random variables.

One can claim that $\hat{p}_i$ and $\hat{r}_i$ are uncorrelated. If so, it can be shown (*26*, p. 29) that

$$\text{Var}\{\hat{m}_i\} \cong (\text{Var}\{\hat{r}_i\}/p_i^2) + r_i^2 \text{ Var}\{\hat{p}_i\}/p_i^4 \tag{3}$$

The first term on the right-hand side is as in Equation 2; the second term accounts for the contribution of $\text{Var}\{\hat{p}_i\}$.

To clarify the meaning of Equation 3, we continue the numerical example from above, in which six injury accidents were reported, constituting some 70 percent of such accidents occurring. We now admit that we are unsure about the 70 percent; perhaps the range 50 to 90 percent covers two standard deviations. This translates into $\text{Var}\{\hat{p}_i\} = 0.04$. Now, an estimate of $\text{Var}\{\hat{m}_i\}$ is $6/0.49 + 6^2 \times 0.04/0.7^4 = 12.2 + 6.0 = 18.2$ accidents^2.

It may be useful to rewrite Equation 3 in order to clarify the meaning of its components. Assume, as before, that we have n annual counts of reported accidents and that both r_i and m_i are expected numbers per annum. Again we make use of the relationships $r_i = m_i p_i$ and $\text{Var}\{\hat{r}_i\} = m_i p_i / n$. Substitution into Equation 3 leads to

$$\text{Var}\{\hat{m}_i\} = m_i/(np_i) + m_i^2 \text{ Var}\{\hat{p}_i\}/p_i^2 \tag{4}$$

It is now clear that the number of years for which accident counts are available (n) affects one component of the variance of $\hat{m}$ but not the other. For no matter how many years we count accidents reported to the police, the uncertainty surrounding $\hat{p}_i$ puts a limit on how accurately m_i can be estimated. We illustrate this point in Figure 3, which was constructed for the case in which 3 years of accident data are available; m is 10 accidents per year and $\text{Var}\{\hat{p}\}$ is 0.04. The ordinate is shown to be constituted of two parts corresponding to the two summands in Equations 3 and 4. The first striking observation is that, in the case shown here, the uncertainty about the proportion of accidents reported ("second term") has a larger effect on the accuracy of estimation than has the randomness in the count of accidents ("first term"). The next feature to note is that when the proportion of reported accidents is small, the second term is so large as to render any attempt at estimation very inaccurate. This is so because the second term is proportional to the reciprocal value of p^2.

Estimation of Relative Changes in $\langle m \rangle$

Often one is not interested in the magnitude of the m_i's or the absolute changes therein. Rather, the goal is to obtain

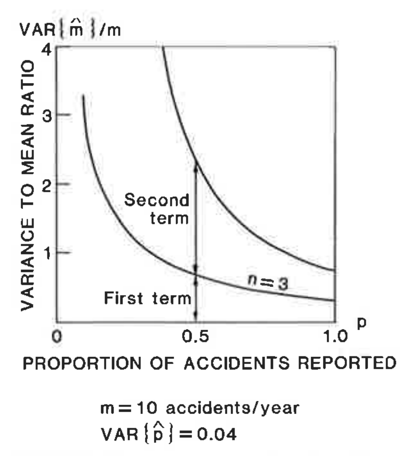

FIGURE 3 Accuracy of estimation when p is not known exactly.

estimates of the *ratios* of the m_i's "before" (or "without") treatment to the m_i's "after" (or "with") treatment. For brevity we denote these ratios as θ_i, $i = 1,2,\ldots$. Estimates of the θ_i's are usually obtained by dividing the number of reported accidents of class i before (without) treatment by the number of accidents in the same class reported after (with) treatment. If p_i is the same for the before and after periods (or for the entities with and without treatment that are being compared), the net effect of incomplete accident reporting is merely to reduce the amount of information that can be collected per unit of time. If, say, only 50 percent of the accidents are reported, it will take twice as long to collect the data that would be available with complete reporting. This is the essence of the relationship in Figure 2.

Unfortunately, even if estimates of the relative change in safety are carefully calculated for several accident classes, difficult problems remain. First, official accident reports contain only coarsely graded estimates of severity, which cannot be very reliable. Second, the probability of an accident's being reported increases not only with accident severity but also with the number of vehicles involved and the age of the persons affected. Therefore, classification by severity alone is obviously insufficient to ensure that the p_i's are the same for the entities whose accident histories are being compared. Third, we know of no empirical evidence to support the assumption that the probability of an accident's being reported is the same in different parts of the same city (when entities with and without treatment are being compared) or that it remains constant during different periods of time for a specific set of entities (as in before and after comparisons) or, in the worst case, that it is the same when one compares the reported accident histories of different entities and different time periods.

It is for these reasons that one must admit the possibility that the p_i's for the entities being compared are not identical. The quantitative implications of such an admission are explored below.

We imagine that the two p_i's, the probabilities that an accident of class i will be reported before (with) and after (without), are drawn at random from a distribution of p_i's that has a mean $E\{p_i\}$ and a variance $\mathrm{Var}\{p_i\}$. To keep the algebra manageable, one has also to invoke the assumption that the probability of reporting and the safety effect of a treatment or measure are statistically independent. With this (possibly questionable) assumption, it can be shown that

$$\mathrm{Var}\{\hat{\theta}_i\} \cong \theta_i^2\ [1/(r_i\theta_i) + 1/r_i + 2\ \mathrm{Var}\{p_i\}/p_i^2] \qquad (5)$$

A numerical example may again be of use. Suppose that the number of reported injury accidents changed from 25 before some treatment to 20 after the treatment. It follows that $\theta_{\text{injury}} = 20/25 = 0.8$, indicating a 20 percent reduction in injury accidents. If the probability of reporting an injury accident were constant during the before and after periods, an estimate of $\mathrm{Var}\{\hat{\theta}_{\text{injury}}\} = 0.8^2[(1/20) + (1/25)] = 0.64(0.05 + 0.04) = 0.058$. However, if an estimate of p_{injury} is 0.7 and an estimate of $\mathrm{Var}\{p_{\text{injury}}\}$ is 0.01, we have to add to the sum in square brackets $2 \times 0.01/0.7^2 = 0.04$. Now, $\mathrm{Var}\{\hat{\theta}_{\text{injury}}\} = 0.083$. In this case, the uncertainty in accident reporting contributes to the variance of the estimate of θ_{injury} an amount that is similar to that contributed by the randomness in accident occurrence before or after treatment.

Equation 5 is similar in nature to Equation 4. It shows that even the accuracy of relative comparisons of safety is severely affected by the variability in the probability of accidents' being reported. This variability imposes a limit on estimation accuracy that is not affected by the amount of data collected. Furthermore, the inaccuracy of estimation grows explosively as the proportion of accidents reported diminishes.

The principal results of this analysis can be summarized as follows. The level of accident reporting plays a central role in the estimation of road safety. Were it realistic to assume that the proportion of accidents reported is constant during a fairly long period of time, a lesser level of reporting would merely correspondingly prolong the time required to collect a fixed amount of accident data. However, it appears unrealistic to so assume. If we admit that the probability of reporting an accident is not known with precision, our ability to estimate road safety is seriously affected. There is now an added term in the variance of $\hat{m}_i$ or $\hat{\theta}_i$. This "second term" increases in direct proportion to the variance of p_i (or its estimate) and with the square of $1/p_i$.

DISCUSSION AND SUMMARY

We set out to describe the extent of accident underreporting and some of its quantitative implications.

It appears that accident underreporting is rather substantial. Estimates of underreporting culled from the literature differ widely. It is therefore difficult to know by how much accidents are underreported in some specific region during a certain period of time. As a ballpark figure, fatalities seem to be known to an accuracy of ±5 percent; perhaps 20 percent of injuries that require hospitalization do not show up in police records; of all injuries sustained in motor vehicle accidents, perhaps half are not reported to the police and the reporting of PDO accidents is likely to be even lower. The probability of an accident's being reported depends on the severity of the outcome, the age of the victim, his or her role in the accident, the number of vehicles involved, and other factors.

The importance of incomplete accident reporting derives from the uses to which the information is put. The use considered here is that of measuring road safety. In this respect there appears to be a curious inconsistency in that road safety must be measured in terms of accidents that *occur*, yet most statements about road safety rely solely on accidents that are *reported*.

The analysis leads to several observations. First, it should be self-evident that one cannot make statements about the size of a road safety problem (the magnitude m for accidents of kind i) without using an estimate of the proportion of accidents in that class that get reported. Equation 1 only reaffirms the obvious. Still, it is rare to encounter statements about the magnitude of a safety problem in which the obvious and the self-evident are given consideration. Most of what is said about safety is confined to statements based on accidents that have been reported and not on estimates of what has occurred. Not only do such statements make the safety problem appear to be smaller than it really is, they also mix and confuse changes and trends in safety with changes and trends in the inclination to report or record accidents.

Second, statistical statements about the accuracy with which the magnitude of m_i's is estimated have to recognize the effect of p_i and Var$\{\hat{p}_i\}$. If this is not done, estimates appear to be more accurate than they really are. The implication is that virtually all that has been said (in safety) about the statistical significance of the differences between means, about the significance of a deviation from an expected value, or about the size of confidence intervals requires reexamination.

Third, accuracy of estimation improves sharply as p_i increases and as Var$\{\hat{p}_i\}$ decreases. This serves to emphasize the importance of completeness in accident reporting by police and supports the investment of effort in the merging of police, hospital, and insurance records. It also supports the need for better information about the magnitude of p_i by accident type.

Fourth, if one is interested only in the ratio of m_i's, that is, in the relative change of safety, the requirement of paramount importance is that the entities for which reported accident histories are being compared ("before" versus "after" or "without" versus "with") have the same probability of an accident's being reported. If that requirement is satisfied, the effect of incomplete reporting is merely that of prolonging the time required to collect a fixed amount of accident data.

Inasmuch as the probability of an accident's being reported is known to increase with its severity and because all real treatments affect both the frequency and the severity of accidents, it seems never correct to estimate the safety effect of some treatment using the ratio of all accidents without to all accidents with treatment. Doing so violates the aforementioned requirement. When estimating the safety effect of some treatment, accidents should be divided into classes in such a manner that, within each class, average severity "without" is likely to be the same as the average severity "with." However, accident severity is not the only factor that influences p. Because p increases with the number of vehicles in the collision, single-vehicle, two-vehicle, and multivehicle accidents should also form separate classes. It is easy to see that the aforementioned requirement is not easy to satisfy in practice.

Fifth, although convenient and tempting, there is no reason to assume that the probability of an accident's being reported is the same for the two sets of entities whose accident histories are being compared. The real question is not whether the two probabilities are equal but how large the difference is between them. The implications of this are captured by Equation 5. Once again, the effect of uncertainty about the difference in the two probabilities is to decrease the accuracy with which the safety effect of the treatment can be known. As the amount of accident data increases, the uncertainty about the differences in the probability of reporting begins to dominate the accuracy of estimation. In fact, it determines a limit on the accuracy with which the safety effect of a treatment can be known irrespective of the amount of available accident data.

In summary, the ability to make quantitative statements about the number of accidents occurring requires that we know what proportion of these accidents gets into police records. Thus, research is needed to find out what the p_i's are by type of accident, by region, by time, and so on. Furthermore, the accuracy of our statements deteriorates rapidly as the proportion of accidents reported diminishes and the uncertainty about the prevailing level of accident reporting increases. Thus, for credible statements about road safety and its changes, the proportion of accidents reported to the police needs to be high, stable over time and location, and accurately known.

ACKNOWLEDGMENT

A part of this research was performed with the support of the National Science and Engineering Research Council of Canada. In the statistical analysis the authors benefited from conversation and correspondence with S. Danielsson of the Swedish Road and Traffic Research Institute.

REFERENCES

1. R. N. Smith. The Reporting Level of California State Highway Accidents. *Traffic Engineering Journal*, June 1966.
2. J. Thorson and J. Sande. Hospital Statistics on Road Traffic Accidents. In *Proceedings, Third Triennial Congress*, International Association for Accident and Traffic Medicine, Copenhagen, Denmark, 1969.
3. R. E. Scott and P. S. Carroll. *Accident Data Inaccuracies.* Highway Safety Research Institute, University of Michigan, Ann Arbor, 1971.
4. J. P. Bull and B. J. Roberts. Road Accident Statistics—A Comparison of Police and Hospital Information. *Accident Analysis and Prevention*, Vol. 5, 1973, pp. 45–53.
5. E. G. House, P. F. Waller, and G. Koch. *How Complete Are Driver Records? Analysis Based on Insurance Claim Crashes.* Highway Safety Research Center, University of North Carolina, Chapel Hill, 1974.
6. R. R. Bourbeau and C. Laberge-Nadeau. Road Injuries in Quebec: Nature, Severity and Direct Costs. In *Proceedings, 21st Annual Meeting,* American Association for Accidental Medicine, Arlington, Ill., 1977.
7. Helmets and Crash Odds for North Dakota Cyclists. *Highway Loss Reduction Status Report,* Vol. 14, No. 1, 1979.
8. J. Greenblatt, M.-B. Merrin, D. Morganstein, and S. Schwartz. *National Accident Sampling System Non-Reported Accident Survey.* Report PB85–145712. U.S. Department of Transportation, Nov. 1981.
9. J. V. Stulginskas et al. A Total Population Survey of Traffic Accidents Among Children. In *Proceedings, 27th Annual Meeting,* American Association for Accidental Medicine, Arlington, Ill., 1983.
10. J. P. Nicholl. *The Use of Hospital In-Patient Data in the Analysis of the Injuries Sustained by Road Accident Casual-*

ties. TRRL SP 628. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1983.
11. A. Windels. Epidemiologic und Struktur Todlicher Verkehrsunfalle in Dortmund 1980. Doctoral dissertation. Medical School, University of Münster, 1984.
12. K. Puschel and B. Brinkmann. Zür Dunkelziffer von Verkehrstoten. *Hamburger Artzerblatt*, Vol. 38, 1984, pp. 162–164.
13. T. P. Hutchinson. Medical Statistics on Road Accident Injury in Several Countries. In *Transport Risk Assessment* (S. Yagar, ed.), University of Waterloo Press, Waterloo, Ontario, Canada, 1984.
14. M. W. Mass and S. Harris. Police Recording of Road Accident In-Patients. *Accident Analysis and Prevention*, Vol. 16, No. 3, 1984.
15. H. Hautzinger et al. Accuracy of Official Road Accident Statistics. *Internationales Verkehrswesen*, Vol. 37, No. 4, 1985.
16. J. I. Barancik and D. Fife. Discrepancies in Vehicular Crash Injury Reporting: Northeastern Ohio Trauma Study IV. *Accident Analysis and Prevention*, Vol. 17, No. 2, 1985, pp. 147–154.
17. P. F. Agran and D. E. Dunkle. A Comparison of Reported and Unreported Non-Crash Events. *Accident Analysis and Prevention*, Vol. 17, No. 1, 1985.
18. W. Siegener and W. Lenhart. *Analyse von Unfalldunbelziffern.* BASt Report Project 7536. BASt, Köln, West Germany, 1985.
19. D. Shinar et al. The Validity of Police Reported Accident Data. *Accident Analysis and Prevention*, Vol. 15, No. 3, 1983.
20. W. L. Kelman. An Analysis of Pedestrian Accidents and the Accident Data Retrieval System in Toronto 1972–75. M.Sc. thesis. University of Toronto, Toronto, Ontario, Canada, 1977.
21. A. S. Hakkert and I. Hocherman. *The Reliability of Reporting of Road Accidents with Casualties.* RSC Report 77/17. Road Safety Centre, Technion, Israel, 1978 (in Hebrew).
22. H. H. Keller. Traffic Records in Germany: Traffic Record Systems Forum. Pergamon Press, Elmsford, N.Y., 1985.
23. C. R. Rao. On Discrete Distributions Arising Out of Methods of Ascertainment. In *Classical and Contagious Distributions* (G.P. Patil, ed.), 1965.
24. C. D. Kemp. An Elementary Ambiguity in Accident Theory. *Accident Analysis and Prevention*, Vol. 5, 1973, pp. 371–373.
25. S. Erlander. A Review of Some Statistical Models Used in Automobile Insurance and in Road Accident Studies. *Accident Analysis and Prevention*, Vol. 3, 1971, pp. 45–75.
26. N. L. Johnson and S. Kotz. *Discrete Distributions.* Houghton Mifflin Company, Boston, Mass., 1969.

Effect of Low-Cost Accident Countermeasures on Vehicle Speed and Lateral Placement at Narrow Bridges

Brian L. Bowman and Philip Brinkman

The results of a research effort sponsored by FHWA are summarized, the primary purpose of which was to determine the effectiveness of low-cost countermeasures in reducing accident potential at narrow bridges. The operational-based evaluation was performed by conducting before-and-after analyses of vehicle speed and lateral placement at 18 narrow bridge approach sites. The low-cost countermeasures that were evaluated consisted of combinations of advance warning signs, pavement markings, raised pavement markers, roadside delineators, type 3 object markers, and adhesive delineators. Measurements of vehicle speed and lateral placement were obtained by using the FHWA Traffic Evaluation System. With one exception, the operational-based effectiveness evaluation did not reveal any statistically significant differences at the 10 percent level between the sites before and after the implementation of the countermeasures. The one exception was that these countermeasures significantly reduced speed variation when all vehicle types and time periods were analyzed together. For this analysis category, therefore, the low-cost countermeasures resulted in more uniform driving behavior.

Highway bridges are a necessary part of any roadway system and have always been the subject of specialized engineering efforts. Their construction requires a more sophisticated engineering design analysis and a higher construction cost than that for the roadways connecting them. In the past the primary purpose of the extra effort has been to ensure that the bridge structure would support dynamic design loads without failure. Until relatively recently, bridge width was not a major concern and would often be reduced for economic reasons. The results of this practice are narrow bridges, especially on the rural road system, which pose a threat to all motorists.

Because bridges are typically designed to provide longer service lives than are the connecting roadways, there are many instances in which the roadway is upgraded and, because of cost constraints, the bridge is not. Because of the relatively high cost of bridge widening and construction, some bridges date back to the early 1900s. Bridge abutments and parapets, many of which are unguarded, present dangerous physical obstructions to motorists. The changes in cross-section width between the approach roadway and the narrow bridge result in traffic flow restrictions and present unexpected hazards to motorists. The result is an increase in erratic driving behavior, fixed-object accidents, and vehicle-vehicle accidents.

The optimal solution would be to upgrade all narrow bridges on U.S. roadways. The extreme costs associated with rebuilding all the deficient bridges on our roadway system make the optimal solution, at least on a short-term basis, unrealistic. The result is that highway agencies are implementing countermeasures designed to reduce crash severity and improve motorist information by providing increased advance warning, delineation, and hazard conspicuity. The rationale behind the implementation of these countermeasures is that if it is not possible to physically protect the motorist from hazards, efforts must be made to provide them with sufficient information to protect themselves. However, it is difficult to ascertain from accident-based studies how effective these low-cost countermeasures are in actually increasing motorist safety.

The difficulty in determining the effectiveness of low-cost countermeasures at narrow bridges with accident-based analysis is the low number of accidents per bridge per year and inaccuracies in identifying the exact accident location from report forms and the exact date on which the countermeasures were installed. In this paper, results are presented of a study that was initiated in response to the recognized difficulties in conducting accident-based effectiveness evaluations of low-cost countermeasures at narrow bridge sites.

STUDY SCOPE AND OBJECTIVES

One of the primary purposes of this research study, sponsored by FHWA, was to determine the effectiveness of low-cost countermeasures at narrow bridges. The study concentrated on analyzing only operational data such as vehicle speed and lateral placement for countermeasures installed during the project tenure. Sites selected for project purposes consisted only of two-lane, single-structure, undivided bridges.

B. L. Bowman, Goodell-Grivas, Inc., 17320 W. Eight Mile Road, Southfield, Mich. 48075. P. Brinkman, FHWA, U.S. Department of Transportation, 6300 Georgetown Pike, HSR-30, McLean, Va. 22101.

BACKGROUND

Narrow bridges have been recognized as a highway safety problem for many years. In 1978, NHTSA reported that the severity of bridge-related accidents was roughly twice that of average accidents. Other studies have revealed that as many as 60,000 bridges are deficient in width (*1*).

Studies have shown that bridge accidents result in high severity rates, as emphasized by the accident experience for Virginia and Kentucky (Table 1). These findings indicate that bridge-related accidents are considerably more severe than other accident types and their frequency represents cause for concern.

Other researchers have also noted the safety problems with bridges. Kaiser determined that traffic accidents at bridges account for twice as many fatalities as do railroad crossing accidents and represent about 3 percent of all accidents in Ohio (*2*). Hilton estimated that narrow bridges account for 1.6 percent of all accidents and 3.4 percent of all fatalities on Interstate highways (*3*).

One of the major reasons that bridges may be hazardous is that many are functionally obsolete, having been built before the adoption of current design standards. Michie states that, on the basis of length alone, a bridge is "more hazardous than the roadway in general and a large number of bridge accidents can be attributed to narrow bridges, obsolete approach guardrails and inadequate bridge rail installations" (*4*). FHWA's national bridge inventory conducted in 1975 shows that 75 percent of the nation's 564,000 bridges were built before 1935 (*5*). In this inventory 20 percent, or 105,000 bridges, is structurally deficient or functionally obsolete, and this number is expected to increase by 2,000 per year. From this national inventory, Weaver and Woods estimated that the number of narrow bridges on two-lane rural roads was 37,000 (*6*).

Although bridge widening is thought to be the most desirable treatment for problems with narrow bridges, its high cost makes it infeasible in most instances (*7*). Mak and Calcote have pointed out that limited resources necessitate the selection of cost-effective treatments, such as signing, roadway delineation, and longitudinal markings (*8*). However, the effectiveness of these countermeasures is uncertain.

DEFINITION OF NARROW BRIDGE

There is no agreed-on single definition of what constitutes a narrow bridge. Most authors agree that width alone cannot be used. AASHTO considers narrow any bridge that has a width less than that of the approach traveled way (*9*). AASHTO also states that the term "narrow" is subjective and should be based on the following characteristics:

Geometrics	*Traffic Characteristics*
Approach roadway width	Approach speed
Approach sight distance	Traffic volume
Bridge width	Percent commercial vehicles
Bridge length	
Horizontal alignment	
Vertical alignment	

Other important factors requiring consideration may include area type and highway functional class. AASHTO provides a table that can be used to classify bridges as narrow on the basis of functional road type, average daily traffic, and percentage of commercial vehicles (*9*, pp. 84–85).

AREA OF BRIDGE INFLUENCE

Narrow bridges can cause accidents that do not occur at or on the physical structure of the bridge itself. Previous research has recognized that driver modifications of behavior on bridge approaches result in changes in vehicle lateral placement and speed, which can result in increased accidents. This requires an appropriate area of influence, which includes roadway segments approaching and leaving the bridge.

Turner and Rowan considered accidents occurring between 1972 and 1979 on 960 bridges on state routes in Alabama (*10*). Accidents on bridge approaches occurred at a rate more than twice as high as the rate as on the adjacent roadway. This increase was found to extend approximately 0.35 mi (573 m) from the bridge ends. Also, police officers were found to record bridge accidents to the nearest 0.1 mi (164 m) in more than half the cases, although some accident report forms required that they be recorded to the nearest 0.01 mi (*10*). This implies that accidents reported as occurring at the center of a short bridge may actually be occurring on the bridge approaches.

In a 1982 study of accidents on narrow bridges, Mak and Calcote collected bridge-related accidents that were coded as occurring on the bridge or within 500 ft (155 m) of either side of the bridge (*8*). This study established that the area of influence of a 200-ft (62-m) bridge totaled 1,200 ft (372 m).

TABLE 1 PERCENTAGE OF BRIDGE-RELATED ACCIDENTS (*1*)

	Interstate/Parkway Highways		Primary/Secondary Highways	
State	Percentage of All Accidents	Percentage of All Fatalities	Percentage of All Accidents	Percentage of All Fatalities
Virginia	3.2	7.1	1.6	3.4
Kentucky	7.6	17.2	2.9	3.8

SOURCE: NHTSA 1978 compilation of data.

The results of accident-based studies indicate a need to consider the approaches on both sides of the bridge when driver-related operational data are collected and when accident-based countermeasures are evaluated. The length should be a minimum of 0.1 mi (0.16 km) on each side of the bridge to account for inaccuracies in accident reporting and changes in vehicle encroachments and speeds on bridge approaches. For highway agencies whose locational reporting accuracy is low, a length of up to 0.3 mi (0.48 km) on each side of the bridge may be appropriate.

COLLECTION OF OPERATIONAL DATA

Measures of Effectiveness

The primary purpose of installing low-cost countermeasures at narrow bridge sites is to reduce accident frequency. Although the accident rate at a narrow bridge site may be twice as high as the rate on a roadway segment, the total number of accidents at a given bridge site is still small. Unlike an evaluation based on operational measures of effectiveness (MOEs), an accident-based evaluation at an individual location therefore requires an unacceptably long period to accumulate sufficient data for statistical validity. This study concentrated on obtaining operational MOEs for various narrow-bridge countermeasures.

Operational evaluations can be conducted shortly after countermeasure implementation (within 1 or 2 months) and require 2 or fewer days of data collection for both the before and the after time periods. Operational MOEs can be used as interim MOEs before an accident-based evaluation. The inherent assumption behind a non-accident-based evaluation is that a significant change in appropriate MOEs (i.e., lateral placement, encroachment, and speed change) indicates improved safety. It is assumed that if a countermeasure results in a significant reduction in vehicles crossing left of the centerline and other hazardous manuevers, this reduction is synonymous with a reduction in accident potential.

Appropriate operational MOEs were selected by establishing a causal chain of the predominant accident types, probable causes, countermeasures, and safety objectives (Figure 1). Low-cost countermeasures at narrow bridge sites are intended to reduce accidents by altering driving behavior. These intended changes in driver behavior are referred to as intermediate objectives in Figure 1. The MOEs selected to evaluate the low-cost countermeasures are primarily related to vehicle speed and lateral position. The logical relationships between these measures and the intermediate objectives are as follows:

1. Mean speed over all tapeswitch deployments. The low-cost countermeasures provide additional driver information and guidance. These driver inputs may result in changes in average speed through the bridge and bridge approach. The expected direction of this change between the before and after time periods is not, however, readily

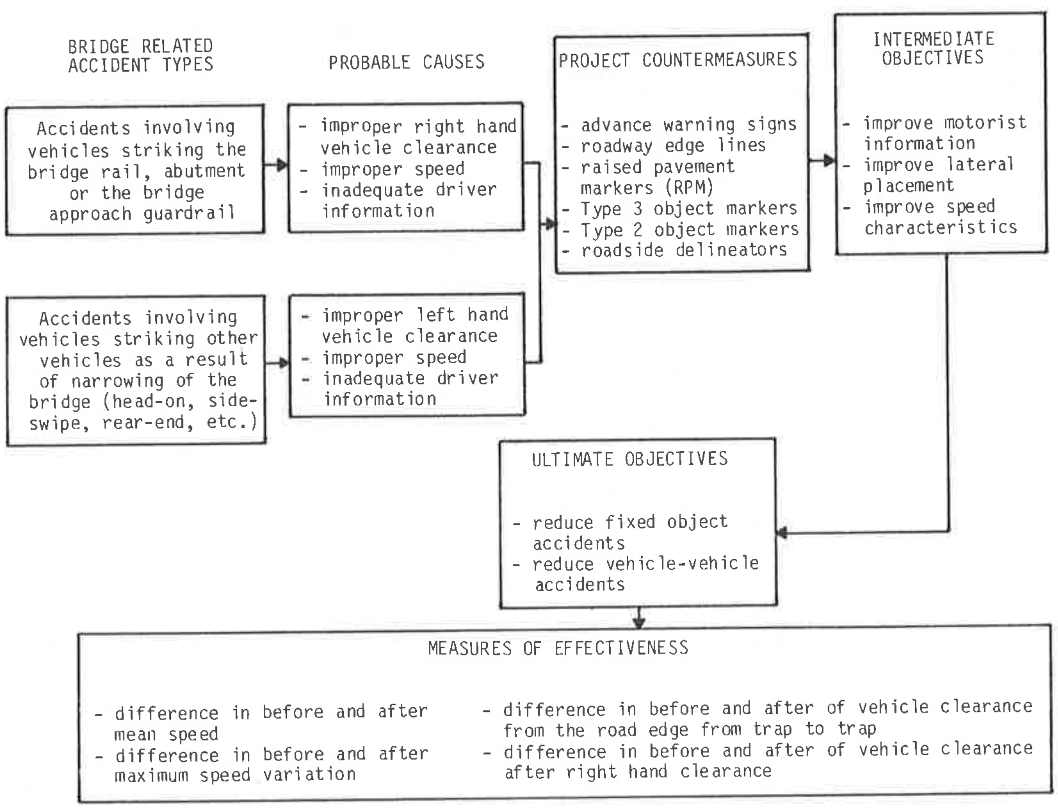

FIGURE 1 Causal chain and appropriate MOEs for low-cost countermeasures at narrow bridge sites.

evident. The installation of a countermeasure to improve driver awareness could result in either an increase or a decrease in average speed, depending on the physical conditions at the bridge site.

For example, some bridge approaches with limited sight distance may pose problems for vehicles approaching too fast and then decelerating rapidly to pass safely over the bridge. In this instance, countermeasures such as advance warning signs would be intended to reduce average speeds on the approach. Evaluation of this countermeasure, using mean speed as the MOE, could interpret a reduction in speed as an indication that the countermeasure is effective. However, consider the case of a narrow bridge where the visibility is so poor (i.e., no lighting or delineation) that motorists must slow down at night to safely traverse the bridge site. An effective delineation treatment (e.g., raised pavement markers, paddle markers, or striping) may improve visibility such that motorists can adequately recognize the bridge site and maintain their approach speed safely. In this instance, an effective countermeasure may result in vehicle speeds that remain unchanged or increase slightly.

2. Maximum speed variation across deployment. This MOE was obtained by measuring the maximum variation in speed exhibited by individual vehicles in the trap array. This maximum speed variation was averaged over all the observations to obtain the analysis value. The increased visual conspicuity and motorist information provided by the countermeasures can logically be expected to result in more uniform speeds through the bridge approach. Speed variability may be indicative of the potential for accidents. A sudden deceleration on the bridge approach could create unexpected hazards, resulting in rear-end (from a trailing vehicle), bridge-related, or head-on accidents (excessive speed causing the inability to maintain proper lateral position). Increased motorist information (adequate delineation or advance warning) could theoretically result in a more gradual deceleration by the motorist and enhanced safety through the bridge site.

3. Mean speed at tapeswitch deployments. This MOE was obtained by averaging the speeds at each trap of every valid vehicle that traversed the test site. The purpose of this MOE was to determine whether the countermeasures resulted in changes to the speed profile at the bridge sites. The most advantageous condition would be to have identical average speeds or a linear reduction in average speed at every tapeswitch deployment. This would be indicative of increased motorist information and confidence in the vehicle guidance tasks required of the narrow bridge site. This analysis differs from the analysis of speed variance in that it provides a measure of the average speed at each trap. The analysis of speed variance used the average speeds at each trap to develop a variance measure of the entire approach site. The analysis of mean speeds by tapeswitch deployment permits the further analysis of which traps had the highest or lowest mean speeds if significant differences are revealed by the statistical analysis of the before and after time periods.

4. Right-hand lateral position at tapeswitch deployments. This measure of lateral placement was selected to provide an indication of the effectiveness of the countermeasure in changing the lateral position of the vehicles. It provides an indication of the potential for accidents with fixed objects located to the right of the roadway and with opposing vehicular traffic. Analyzing lateral position from trap to trap allows determination of the change in right-hand distance from trap to trap and of the location on the approach at which these changes occurred. The lateral placement measures were obtained by measuring the distance from the right edge of the paved roadway surface to the outside edge of the right front tire.

5. Deviations in right-hand lateral placement between tapeswitch deployments. This MOE was obtained by determining the differences in the average right-hand lateral distance between adjacent tapeswitch deployments for both the before and the after time periods. The purpose of these analyses was to determine whether the countermeasures were effective in providing increased motorist guidance, resulting in a more uniform vehicle path.

Many of the low-cost countermeasures evaluated during the project consisted of treatments that would benefit the motorist primarily at night. To evaluate the effect of light conditions on MOE effectiveness, the data were collected separately for day and night conditions. The type of vehicle was also noted, to permit a determination of whether various classes of vehicles are affected differently by the implemented countermeasures.

Characteristics of Selected Test Sites

Nine narrow bridge sites were selected for analysis. Data were obtained from both approaches to the nine bridges, resulting in measurements on 18 approaches. A summary of the physical characteristics of each approach is presented in Table 2, inspection of which shows that all of the narrow bridges selected for analysis were less than 24 ft (7.3 m) in total width (curb to curb). Approach widths were measured as the total distance from roadway edge to roadway edge. The bridge directional width was obtained by measuring from the curb, when present, or from undisturbed debris near the bridge rail [approximately 6 in. (152 mm)] to the center of the centerline on the bridge deck. For all the test sites, total bridge width was less than the approach roadway width.

All the test sites were located in rural environments with one-way volumes that varied from a minimum of 800 to a maximum of 2,625 vehicles per day. The majority of approaches consisted of straight roadway sections with sight distances greater than 900 ft (279 m). Those locations that had reduced sight distances because of horizontal and vertical curves were posted at speeds below 55 mi/hr (88 km). In all cases, the available sight distance was greater than the minimum safe stopping sight distance recommendations of AASHTO for the posted speeds (*9*, p. 138).

TABLE 2 PHYSICAL FEATURES AT NARROW-BRIDGE TEST SITES

Site and Approach	Total Bridge Width (ft)	Length (ft)	Percent Reduction		Alignment and Sight Distance
			Roadway to Bridge	Roadway and Shoulder to Bridge	
11	24.0	24	13.6	35.4	Straight +900'
12			14.5	36.1	Vertical Curve 600'
21	18.4	50	8.0	34.3	Straight +900'
22			10.7	18.6	Straight +900'
31	20.5	56.6	5.1	28.3	Vertical Curve 600'
32			9.7	26.0	Straight +900'
41	19.9	39.5	16.0	40.9	Horizontal Curve 321'
42			17.1	23.5	Horizontal Curve 525'
51	18.0	44	18.6	35.9	Straight +900'
52			20.0	36.8	Straight +900'
61	20.2	46	11.0	29.6	Straight +900'
62			9.4	28.6	Straight +900'
71	20.3	82.2	9.0	21.0	Straight +900'
72			9.4	21.3	Straight +900'
81	19.4	44	20.5	40.1	Vertical and horizontal curve +900'
82			21.8	40.9	Vertical and Horizontal curve +900'
91	18.5	42	1.6	31.0	Straight +900'
92			2.6	31.5	Straight +900'

Descriptions of Implemented Countermeasures

The selection of countermeasures for project implementation was based on a consideration of what was already present and the standard practice of the respective highway agency. Standard practice for some agencies, for example, did not include the installation of raised pavement markers. In such instances, raised pavement markers were not considered for installation because they would have resulted in roadway conditions that were abnormal, especially at night, for the driver expectancy of area motorists.

A summary of the traffic control and delineation devices that were present before countermeasure installation is presented in Table 3. The countermeasures were always installed as additions to existing conditions, the only exception being mutually exclusive countermeasures such as different edgeline widths. One test bridge, for example, initially had two Type 3 object markers; an additional four were added as part of the project countermeasures. There were therefore a total of six Type 3 object markers, three on each side of the approach, present during the after time period.

Collection of Field Data

Field data were collected by using FHWA's fully automated Traffic Evaluation System (TES). The TES is a computerized data collection system that receives input through a series of tapeswitches. The tapeswitches consist of two copper strips separated by a thin plastic divider along each edge of the switch. As a vehicle passes over the switch, its weight causes the strips to come in contact, which closes the circuit. The electrical impulse generated by each closed circuit is transmitted to a rheostat, which identifies the switch location, and the resultant current triggers the recording of the time, the switch code, and the location code.

TABLE 3 TRAFFIC CONTROL FEATURES AT EACH APPROACH TO NARROW-BRIDGE TEST SITES

Traffic Control Feature	Number of Approaches	
	Before	After
Edgeline		
4 inches	16	12
6 inches	2	2
8 inches	0	4
Post Delineators (Type 2)		
Left Hand Side	0	12
Right Hand Side	0	12
Type 3 Object Markers		
2 on Each Approach	18	8
4 on Each Approach	0	10
Raised Pavement Markers		
Left and Right Hand Sides	4	6
Centerline	4	6
Adhesive Delineation Markers	2	6
Narrow Bridge Signs	10	14
Centerlines		
Approach Marking		
Solid	15	15
Skip	3	3
Bridge Marking		
Solid	9	9
Skip	5	5

Four tapeswitch stations, each configured as presented in Figure 2, were deployed on each narrow bridge approach to record the speed, vehicle type, vehicle width, and lateral placement of traffic. The lateral placement was determined by applying the vehicle speed, measured from Traps A and B, in conjunction with the known angle and distance of Tapeswitch C. The approximate positions at which the four tapeswitches were deployed are shown in Figure 3 and described as follows:

- At a free flow point on the narrow bridge approach—An additional diagonal switch was installed at this location to determine vehicle width, which was necessary for the determination of encroachments. The free flow point was determined to exist at a distance from the bridge that was equal to or beyond the safe stopping sight distance.
- At points that were two-thirds and one-third of the safe stopping sight distance.
- At the beginning of the bridge.

The above guidelines were used for the deployment of TES tapeswitches, but the actual deployment was dependent on the physical site characteristics. Roadway surface condition, the location of physical features (such as trees) for securing the TES unit, and other site characteristics resulted in variations of the actual tapeswitch locations.

ANALYSIS OF OPERATIONAL DATA

The analyses of the MOEs related to vehicle placement and speed were performed using a before-and-after experimental design. The before period consisted of TES deployment ahead of the installation of any countermeasures. The after-period data were obtained after the countermeasures had been in place for at least 2 months. The 2-month waiting period was used to allow any possible novelty effects to dissipate before data collection.

The before-and-after design was considered appropriate because (a) data were being collected at the same sites for each time period, (b) the amount of total elapsed time between finishing the before-and-after data collection tasks was less than 4 months, and (c) the total amount of data collection at each site generally exceeded 24 hr. The result was relatively large sample sizes obtained within a short time interval. The possible effects of biasing factors such as trends over time and regression to the mean were not, therefore, considered threats to statistical validity as they often are when accidents are used as MOEs.

The data obtained for both the before and the after time periods were divided into day and night conditions and into categories by vehicle type. The categories of vehicle type were determined by such criteria as the number of axles, wheelbase, and wheelpath width.

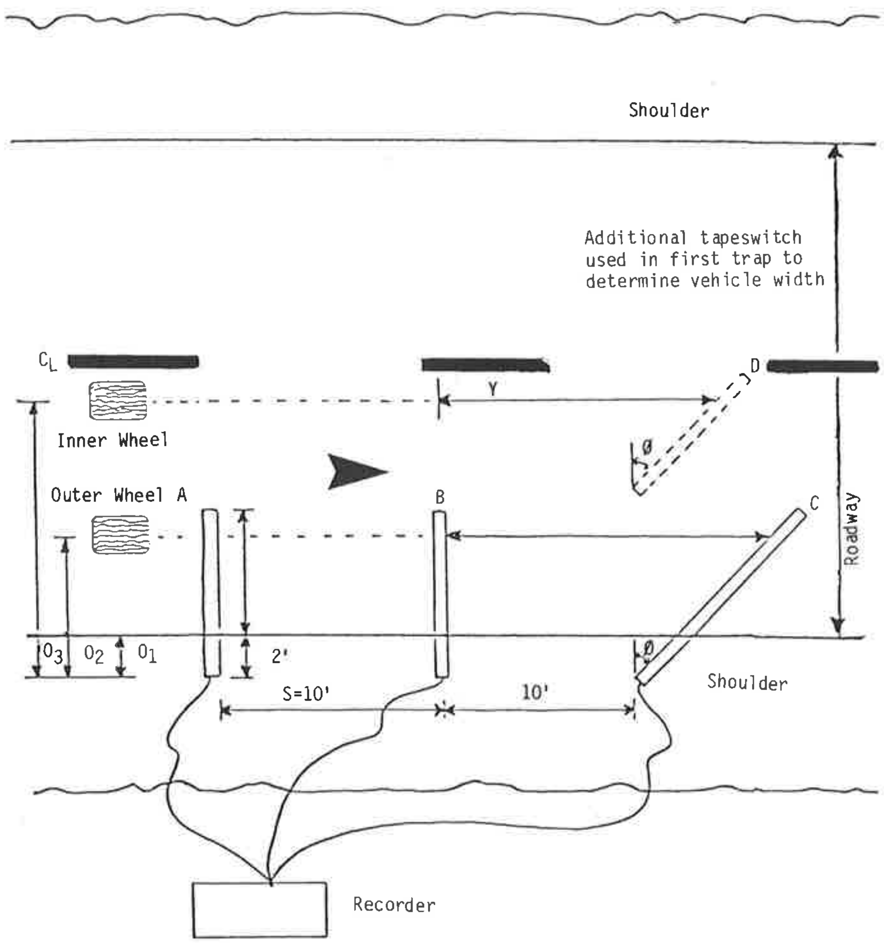

FIGURE 2 Typical tapeswitch deployment.

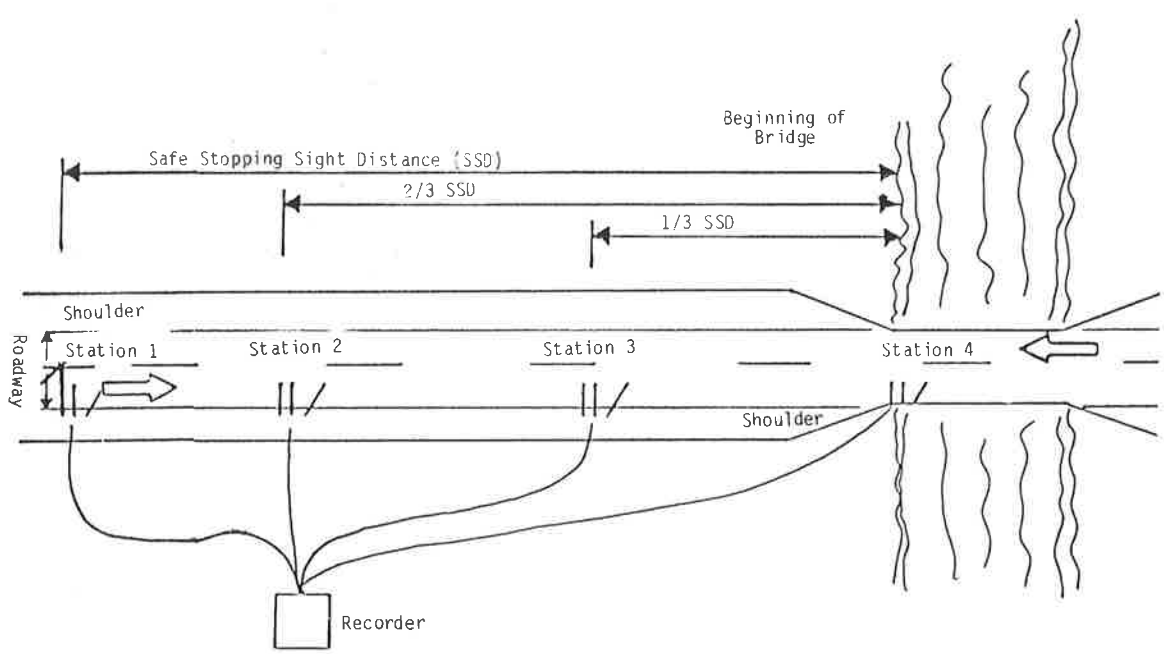

FIGURE 3 Typical approach layout of TES tapeswitches (same tapeswitch configuration set up on other bridge approach).

Type of Analysis

The type of analysis performed was dependent on the MOE and whether each MOE was obtained on a site or a trap-to-trap basis. All the data were divided into categories of bridge approach by time of day and vehicle type. A significance level of 10 percent (i.e., level of confidence of 90 percent) was used for all the statistical tests in this study.

Analyses Based on Individual Observations

The software logic of TES enabled the system to identify a vehicle at the first tapeswitch deployment and to follow that vehicle through the trap array. A unique identifier was assigned to each vehicle, and the speed and lateral position at every tapeswitch deployment were recorded as part of that vehicle's data. It was possible, therefore, to determine the speed and lateral position changes exhibited by each vehicle as it progressed through the trap array. The individual vehicle data were used as input for the first battery of statistical tests. The resulting group means and statistical tests were thus based on large sample sizes and degrees of freedom. Statistical analyses between before-and-after MOE values of individual vehicle measures were conducted by using computerized statistical analysis packages (*11*).

Analyses Based on Individual Vehicle t-*Test Results*

The results of the statistical analysis of individual vehicle measurements consisted of determinations as to whether individual sites or individual traps exhibited a significant difference in their before-and-after MOE values. The sign test was then applied to the results of the site-specific analysis. The purpose of the sign test was to determine whether there were a sufficient number of instances in which a significant difference existed to conclude that the countermeasures resulted in a net effect.

Analysis of Mean MOE Values

The mean MOE values generated by the *t*-tests on the individual vehicle data were analyzed by site and by tapeswitch deployment. This was accomplished by using the appropriate mean from each test site and performing statistical tests to determine whether significant changes had occurred between the before and after time periods.

Mean Speed Over All Tapeswitch Deployments

The before and after mean speeds across all of the tapeswitch deployments were first analyzed by applying the *t*-test to determine whether the magnitude of any exhibited changes was sufficiently large to be considered statistically significant at the 10 percent level within each test site. The purpose of this test was to determine whether the countermeasures installed at a particular site had an impact on the mean speed at that site.

Inspection of the direction of significant changes in mean speed did not reveal discernible patterns. Only three sites had an increase with no accompanying decrease in mean speed across all analysis categories (i.e., categories of vehicle types and time of day). Each of these test approaches consisted of straight roadway sections with no sight restrictions. The test approaches with horizontal curves displayed a greater number of significant mean speed reductions across analysis categories than speed increases (eight reductions versus four increases).

The results of the *t*-test were also investigated to determine whether a sufficient number of increases or decreases in mean speed had occurred within the test sites to signify the presence of trends. This was accomplished by applying the sign test to determine whether, at a 90 percent level of confidence, an increase or decrease in mean speeds could be expected to occur from the installation of low-cost countermeasures. The results of these tests, shown in Table 4, indicate that the probability of observing an equal or greater number of changes due to chance alone exceeds the desired significance level of 0.10. It cannot be concluded, therefore, that the countermeasures being evaluated were the cause of the observed speed changes.

The mean speeds at each test site were also analyzed to determine whether there were any statistically significant differences between the before and after data among the test sites. These analyses were performed separately for day and night conditions in the category of all vehicle types. The results of these analyses are given in Table 5, which reveals that the probability for neither the day nor the night conditions indicated a statistically significant difference at the 0.10 significance level. The low-cost countermeasures did not, therefore, result in significant changes in the mean speeds when evaluated among all the test sites.

Speed Variation Across Deployment

Analysis of the maximum speed variation within sites was performed by applying the *t*-test. This analysis indicated that only three approach test sites experienced a significant increase in speed variability at the 10 percent level. All of the remaining approach sites experienced either a reduction in speed variability across all analysis categories or no significant change.

The speed variations were analyzed by the sign test to determine whether the frequencies with which increases and decreases in variability occurred between the sites were significantly different. Inspection of the summary of the sign test (Table 6) reveals that there is a significant reduction in speed variability for the category of all vehicle types when analyzed for all time periods. The low-cost countermeasures do, therefore, result in more uniform driving behavior. This uniformity was not evident, however, when the data were analyzed separately for day and night conditions.

TABLE 4 SIGN TEST ON CHANGE IN MEAN SPEED BETWEEN BEFORE AND AFTER TIME PERIODS

	All time periods and vehicle types	All vehicle types	
		Day	Night
significant increases (+)	8	6	5
significant decreases (-)	4	3	4
no significant difference	6	9	9
probability of a greater number of speed increases	0.19	0.25	0.50

NOTE: Data are in miles per hour (1 mph = 1.6 km/hr).

TABLE 5 STUDENT'S *t*-ANALYSIS OF MEAN SPEEDS AT EACH TEST SITE FOR ALL TYPES OF VEHICLES

	Day		Night	
	Before	After	Before	After
Mean	49.00	48.95	49.02	48.89
Standard Deviation	4.99	4.59	4.74	4.36
t value	0.03		0.08	
degrees of freedom	34		34	
probability	0.98		0.93	

NOTE: Data are in miles per hour (1 mph = 1.6 km/hr).

TABLE 6 SIGN TEST ON DIRECTION OF CHANGE FOR MAXIMUM SPEED VARIATION

	All Time Periods and Vehicle Types	All Vehicle Types	
		Day	Night
significant increase (+)	2	3	2
significant decrease (-)	9	7	6
no significant difference	7	8	10
probability of a greater number of speed increases	0.03[a]	0.17	0.15

NOTE: Data are in miles per hour (1 mph = 1.6 km/hr).
[a]Denotes significant difference at level of 10 percent.

Student's *t*-test was performed on the mean speed variation to determine whether there were statistically significant differences between the before and after data for day and night conditions. A summary of the analysis is given in Table 7. Note that the overall means for all of the time periods are relatively close to each other. Because the Student's *t*-test probabilities are greater than 0.10, it cannot be concluded that a significant difference exists between the before and after measurements for either day or night conditions.

Mean Speed at Tapeswitch Deployments

Vehicle speeds at each tapeswitch deployment were inspected to determine whether the speed profile of motorists changed because of the installation of the low-cost countermeasures. This analysis indicated that trends that were present during the before time period continued into the after period. For example, four site approaches exhibited higher speeds at the bridge during the before period than at any other tapeswitch locations on their respective roadway approaches. This trend continued into the after period.

A paired t-analysis was performed on the trap data to ascertain whether there were significant differences in the before and after time periods. This analysis was performed by considering the data from the different time periods for each trap as paired observations. The paired t-analysis, for example, resulted in pairing the before data from Trap 1, approach 12, with the after data from Trap 1, approach 12. The paired t-analysis compensated for the differences in trap distance from the bridges. The results of the paired t-analysis, performed separately for day and night conditions on the category of all vehicle types, are summarized in Table 8. At a significance level of 10 percent, no statistically significant differences were indicated by either the day or night data sets. It cannot be concluded at a 90 percent level of confidence that the low-cost countermeasures resulted in significant changes in speed between tapeswitch locations.

Lateral Position at Tapeswitch Deployments

The analyses of lateral position at tapeswitch deployments were conducted to determine whether the countermeasures caused lateral position variations within each site and where on the approach these variations occurred. Lateral placement was determined by measuring the distance from the right edge of the paved roadway surface to the outside edge of the right front tire.

This inspection revealed that 10 and 6 approach sites experienced average movements to the right and left, respectively, at the traps closest to the bridge after countermeasure installation. Because of the way in which the directional movements are distributed among the approach sites, it is difficult to associate the direction of movement with the type of countermeasure installed. At four approach sites, 8-in. (203-mm) edgelines were installed as part of the countermeasure. At two of these approach sites, average movements to the left were experienced; at one site, there were movements to the right; and at the remaining site there was no change in any direction. There is therefore no evident direction of movement that can be associated with a countermeasure type.

Paired t-analyses were performed on the tapeswitch deployments to ascertain whether the differences experienced at each tapeswitch deployment were sufficiently large to be significant. These analyses were performed by considering the data from different time periods for each trap as paired observations. The paired t-analyses compensated for the differences in trap distance from the bridges. The results of the paired t-analysis, performed separately for day and night conditions on the category of vehicle type, are summarized in Table 9. No significant sign differences were indicated at the 10 percent level for either the day or the night data. It cannot be concluded, therefore, at a 90 percent level of confidence that the countermeasures resulted in significant changes in lateral placement between tapeswitch deployments.

Deviations in Lateral Placement Between Tapeswitch Deployments

Analyses were performed on the average variation that occurred between adjacent tapeswitch deployments and

TABLE 7 ANALYSIS OF MEAN VARIATION IN SPEED FOR ALL VEHICLE TYPES

	Day		Night	
	Before	After	Before	After
Mean	4.29	4.09	4.15	3.90
Standard Deviation	1.30	1.29	1.27	1.29
t value	0.48		0.58	
degrees of freedom	34		34	
probability	0.64		0.56	

NOTE: Data are in miles per hour (1 mph = 1.6 km/hr).

TABLE 8 PAIRED t-ANALYSIS OF MEAN SPEEDS AT TAPESWITCH DEPLOYMENTS

	Day		Night	
	Before	After	Before	After
Mean standard deviation	48.50 5.13	48.61 4.61	48.48 4.64	48.59 4.56
t Value degrees of freedom probability	-0.80 71 0.43		-0.59 71 0.56	

NOTE: Data are in miles per hour (1 mph = 1.6 km/hr).

between deployments that were the farthest apart. The purpose of these analyses was to determine whether the countermeasures were effective in providing increased motorist guidance resulting in a more uniform vehicle path. The data for these analyses were obtained by determining the difference in the lateral placement for the appropriate trap pairs.

Inspection of the resultant differences revealed that the type of movements between adjacent trap pairs remained relatively constant in both the before and after time periods. Those pairs that exhibited average movements to the right between the traps in the before period usually exhibited movements to the right in the after period. This observation was supported by the results of the paired *t*-analyses that are summarized in Table 10. There were no significant differences at the 10 percent level between the lateral movements exhibited by adjacent pairs in the before or after time periods.

Paired *t*-analyses performed on the differences in lateral movement between the farthest trap pairs (i.e., tapeswitch deployments 1 and 4) are summarized in Table 11. This analysis did not display any significant differences at the 10 percent level between the lateral movements of the farthest trap pairs in the before and after time periods.

CONCLUSIONS

The conclusions presented below are based on the results of the analysis of field data and the literature review.

1. Analysis of individual vehicle speeds indicated that the effects of the low-cost countermeasures were essentially the same for day and night conditions. Three sites during the day and four during the night experienced a significant decrease in speed. When both day and night conditions

TABLE 9 PAIRED *t*-ANALYSIS OF LATERAL POSITION OF TAPESWITCH DEPLOYMENT

	Day		Night	
	Before	After	Before	After
mean	3.71	3.69	3.95	3.97
standard deviation	0.60	0.67	0.60	0.67
t value	0.38		-0.58	
degrees of freedom	70		71	
probability	0.71		0.56	

NOTE: Data are in feet (1 ft = 0.31 m).

TABLE 10 PAIRED *t*-ANALYSIS OF LATERAL POSITION CHANGE BETWEEN ADJACENT TAPESWITCH DEPLOYMENTS

	Day		Night	
	Before	After	Before	After
mean	-0.06	-0.07	-0.13	0.69
standard deviation	0.68	0.77	-0.13	0.78
t value	0.16		-0.01	
degrees of freedom	51		53	
probability	0.87		0.99	

NOTE: Data are in feet (1 ft = 0.31 m).

TABLE 11 PAIRED *t*-ANALYSIS OF OVERALL DIFFERENCE IN LATERAL PLACEMENT

	Day		Night	
	Before	After	Before	After
Mean	-0.20	-0.25	-0.40	-0.40
Standard Deviation	0.92	0.72	0.89	0.72
t value	0.32		-0.02	
degrees of freedom	17		17	
probability	0.76		0.983	

NOTE: Data are in feet (1 ft = 0.31 m).

were analyzed together, eight sites experienced a significant increase, four a significant decrease, and six no significant difference in mean speeds between the before and after time periods. These results did not establish a sufficient difference in the mean speed increases or decreases to attribute the effects to the low-cost countermeasures. The low-cost countermeasures cannot therefore be assumed to result in significant changes in mean speeds.

2. An inspection of the mean individual vehicle speed at each tapeswitch deployment was performed to determine whether the speed profile of motorists changed because of the installation of the low-cost countermeasures. Inspection of the mean speeds at each trap deployment revealed that trends that were present in the before time period continued into the after period. Those sites that exhibited peak speeds at the trap located closest to the bridge during the before period also exhibited peak speeds at the bridge during the after period. It cannot be concluded at a 10 percent significance level that the low-cost countermeasures resulted in significant changes in mean speed between tapeswitch deployments.

3. Estimates of vehicle lateral placement were obtained by measuring the distance from the right road edge to the outside of the right front tire. Inspecting the manner in which the directional movements were distributed among the approach sites resulted in difficulty associating the direction of movement with the types of countermeasures installed. For example, four approach sites received 8-in. (203-mm) edgelines as part of their physical upgrade. At two of these sites average movements to the right, at one site movements to the left, and at one site no change in any direction were experienced. It could not be concluded that the low-cost countermeasures resulted in statistically significant changes in right-hand lateral placement between tapeswitch deployments.

4. Analyses were performed on the average variation that occurred between adjacent tapeswitch deployments and between deployments that were the farthest apart. The purpose of these analyses was to determine whether the low-cost countermeasures resulted in a more uniform vehicle path. Inspection of the differences indicated that the type of movement between adjacent trap pairs remained relatively constant between the before and after time periods. Those pairs that exhibited average movements to the right between the traps in the before period usually exhibited movements to the right in the after period. There were no significant differences at the 10 percent level between the lateral movements exhibited by either adjacent pairs or the farthest trap pairs in the before and after time periods.

5. Estimates of the maximum speed variation were obtained by measuring the greatest difference in speed exhibited by individual vehicles as they progressed through the trap array. This maximum speed variation was averaged over all the observations to obtain the analysis value. The intuitive logic in the selection of this MOE was that a reduction in speed variation denotes increased safety because of the more uniform speeds. This effect was expected to be more pronounced during the nighttime and periods of low visibility when the delineators, edgelines, and hazard markers provide maximum conspicuity. Analyzing data obtained by combining the day and night observations into one group revealed that a significant number of analysis sites experienced a reduction in speed variability after countermeasure implementation at the 10 percent level of significance. When the average speed variation was analyzed separately for day and night conditions, however, there were no significant differences between the before and after time periods.

6. The inability of the MOEs to exhibit statistically significant differences between the before and after time periods can be interpreted in two ways. The first interpretation is that the operational MOEs related to vehicle speed and position are not appropriate measures for narrow bridge sites. The literature review indicated that these measures had been used successfully in prior studies at narrow bridge sites. The use of TES in this study, however, resulted in much larger data bases and greater accuracy than in those studies that relied primarily on manual data collection techniques. In addition, because the narrow bridges studied existed on low-volume rural roadways, the majority of the roadway users were expected to be local motorists who were familiar with the roadway geometrics. These motorists know the presence of narrow bridges and developed driving patterns to safely negotiate the hazardous roadway feature before the installation of the low-cost countermeasures. Their driving characteristics may not, therefore, have been altered by the installation of low-cost countermeasures.

The second possible interpretation is that the countermeasures are not effective in influencing driver behavior. However, the inability of the operational MOEs to identify changes in driving behavior does not necessarily imply that the low-cost countermeasures are ineffective. Accidents are relatively rare events that result from circumstances related to the driver, vehicle, roadway, or environment, or to more than one of these variables. The low-cost countermeasures provide increased delineation and driver information. The impact of these enhancements on potential accidents involving unfamiliar drivers, impaired drivers, and unfavorable environmental conditions (such as restricted visibility and wet and slippery road conditions) cannot be ascertained by analyzing average operational measures. A determination of the actual effectiveness of low-cost countermeasures, therefore, requires a proper accident-based evaluation.

REFERENCES

1. D. L. Ivey, R. M. Olson, N. E. Walton, and G. D. Weaver. Safety at Narrow Bridge Sites. Presented at the 62nd Annual Meeting, AASHTO, Birmingham, Ala., Nov. 1976.
2. H. Kaiser. Traffic Accidents on Highway Bridges on Rural State Highways in Ohio. In *Proceedings of 10th Annual Ohio*

Engineering Conference, Ohio State University, Columbus, 1956.
3. M. H. Hilton. Some Case Studies of Highway Bridges Involved in Accidents. Presented at the 52nd Annual Meeting of the Transportation Research Board, Washington, D.C., 1973.
4. J. D. Michie. Strategy for Selection of Bridges for Safety Improvement. In *Transportation Research Record 757*, TRB, National Research Council, Washington, D.C., 1980, pp. 17–22.
5. *The National Highway Safety Needs Report*. U.S. Department of Transportation, March 1976.
6. G. D. Weaver and D. L. Woods. *No-Passing Zone Treatments for Special Geometric and Traffic Operational Situations*. Texas Transportation Institute, Texas A&M University System, Austin, Sept. 1981.
7. U.S. Congress. House. Committee on Public Works. Subcommittee on Investigations and Review. *Hearings: Highway Safety Design and Operations: Narrow Bridges—Driver Dilemmas*. 93rd Congress, June 1973.
8. K. K. Mak and L. R. Calcote. *Accident Analysis of Highway Narrow Bridge Sites*, Vol. 2: *Technical Documentation*. FHWA, U.S. Department of Transportation, Feb. 1983.
9. *Highway Design and Operational Practices Related to Highway Safety*, 2nd ed. AASHTO, Washington, D.C., 1974.
10. D. S. Turner and N. J. Rowan. Investigation of Accidents on Alabama Bridge Approaches. In *Transportation Research Record 847*, TRB, National Research Council, Washington, D.C., 1982, pp. 1–6.
11. *SPSS-X*. McGraw-Hill Book Company, New York, N.Y., 1983.

Bayesian Identification of Hazardous Locations

JULIA L. HIGLE AND JAMES M. WITKOWSKI

A Bayesian analysis of accident data is used in the identification of hazardous locations. The Bayesian model used in the analysis is developed and discussed. Empirical comparisons of the results from the Bayesian analysis and from classical statistical analyses are also included. These comparisons suggest that there is an appreciable difference among the various identification techniques and that some classically based statistical techniques may be prone to err in the direction of false negatives.

One problem of ongoing interest in highway safety analysis is the identification of hazardous locations on the basis of historical data. Typically, a site is deemed hazardous if its recent accident history exceeds some specified level. One of the most common methods used in practice is to identify a site as hazardous if its accident rate over some period of time exceeds the mean accident rate over all sites in the region plus a multiple of the standard deviation of the site accident rates within that region over the same period of time. Such methods are based on the concept of confidence intervals within the context of classical statistics. The multiple used depends on the degree of confidence desired. Another commonly used technique is the rate-quality method (*1, 2*), which is based on statistical quality control procedures. This technique is used to calculate a critical accident rate, which depends on the degree of confidence desired, for each location. With the rate-quality method, a site is identified as hazardous if its observed accident rate exceeds its critical rate.

It is commonly acknowledged that because of the random variations that are inherent in accident phenomena, historical accident data do not always reflect long-term accident characteristics accurately. A site with a low accident rate (i.e., in the long run) may still have a high accident rate over a short period of time, and vice versa. Thus, the identification of hazardous locations is an inexact science at best. Regardless of the identification method used, traffic analysts will generally agree that the accident rate associated with a particular site is a random variable, a quantity that cannot be predicted with absolute certainty. Moreover, although regional accident characteristics may provide some useful information regarding the accident rate at a particular site, each site must be evaluated separately and should only be compared with sites that have similar underlying characteristics. The vast differences in accident histories that one finds among various sites suggest that the random variables used to describe the accident rates should differ from site to site.

To overcome some of the difficulties associated with the identification of hazardous locations, researchers have increasingly advocated the use of Bayesian analysis in this identification process (*3–7*). Bayesian analysis provides a framework wherein regional accident characteristics can be combined with site-specific accident histories, which results in a coherent method by which the random variables representing the accident rates at the various sites can be mathematically defined. Moreover, by using a Bayesian identification technique, one can identify hazardous sites on the basis of the probability that the accident rate exceeds some level. Such probabilistic identification methods differ both qualitatively and quantitatively from the confidence-based identification methods.

The research reported in this paper can be viewed as a complement to the research presented by Hauer and Persaud (*4–7*), although the techniques used differ substantially. These earlier papers are concerned with predicting the number of accidents that will occur at a particular location, and our research revolves around the accident rate at a particular location. An accurate prediction of the number of accidents at a particular site is invaluable in the assessment of the effectiveness of an improvement program, especially when one considers the phenomenon of regression to the mean. However, before an improvement program is implemented, one must first decide which sites require improvement. The contribution of this paper lies in the identification phase of the improvement process. Specifically, we develop a method for identifying hazardous locations on the basis of a Bayesian analysis of the accident data.

In this paper, we present the results of a Bayesian analysis of accident data from the jurisdiction of the Pima County Department of Transportation in Tucson, Arizona. The paper is divided into a discussion of the Bayesian methodology used in the study, a description of the data used, a comparison of the results of our Bayesian analysis with the results of an analysis based on classical statistical techniques, and our conclusions.

J.L. Higle, Department of Systems and Industrial Engineering, University of Arizona, Tucson, Ariz. 85721. J.M. Witkowski, JHK and Associates, 111 S. Church Street, Suite 470, Tucson, Ariz. 85701.

BAYESIAN METHODOLOGY

Bayesian analysis differs significantly from the classical statistical analysis of accident data. The motivation for the use of the Bayesian analysis is the desire to treat the actual accident rate (i.e., the number of accidents per million vehicles entering an intersection) at a particular location as a random variable and to use a combination of the regional accident characteristics and the accident history at that location to determine the probability that the location is hazardous. In this way, we hope to better utilize the available information throughout the identification process.

Our Bayesian analysis uses a two-step procedure. In the first step, we aggregate the accident histories across a number of sites (i.e., across all sites within an appropriately defined region). The result of this step is a gross estimation of the probability distribution of the accident rates across the region. We then use this regional distribution and the accident history at a particular site to obtain a refined estimation of the probability distribution associated with the accident rate at that particular site. Naturally, we obtain this refined estimation for all sites within the region, and two sites with equivalent histories will have identically refined distributions. This essentially concludes the Bayesian portion of the analysis. With the collection of refined distributions, one can now assess the probability that any given site is hazardous.

To formally describe the Bayesian identification process, we require the following notation:

$\tilde{\lambda}_i$ = accident rate at location i (note that $\tilde{\lambda}_i$ is treated as a random variable);
N_i = number of accidents at location i during the period of time in question;
V_i = number of vehicles passing through location i during the period of time in question;
$f_i(\lambda \mid N_i, V_i)$ = probability density function associated with the accident rate at location i, given the observations N_i and V_i; and
$f_R(\lambda)$ = probability density function associated with the accident rate across the region.

Thus, $f_R(\lambda)$ represents the gross estimation of the probability distribution of the accident rate across the region, and $f_i(\lambda \mid N_i, V_i)$ represents the refined estimation of the probability distribution at site i, as previously discussed. Moreover, the cumulative distribution function associated with the accident rate, $\tilde{\lambda}_i$, is given by

$$P\{\tilde{\lambda}_i \leq \hat{\lambda}\} = \int_0^{\hat{\lambda}} f_i(\lambda \mid N_i, V_i)\, d\lambda$$

In performing this analysis, we make the following assumptions, which are similar to those of Morin, Norden et al., Hauer and Persaud, and Glauz et al. (*1, 2, 4, 6–8*), to name but a few.

A1. At any given location, when the accident rate is known (i.e., if $\tilde{\lambda}_i = \lambda$), the actual number of accidents follows a Poisson distribution with expected value λV_i. That is,

$$P\{N_i = n \mid \tilde{\lambda}_i = \lambda,\ V_i\} = \frac{(\lambda V_i)^n}{n!} e^{-\lambda V_i}$$

A2. The probability distribution of the regional accident rate, $f_R(\lambda)$, is the gamma distribution.

The first assumption indicates that because the actual accident rate is explicitly treated as a random variable, the conditional distribution of the number of accidents (given the accident rate) is the Poisson distribution. The second assumption implies that

$$f_R(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

for some α and β. Thus, the first step associated with the Bayesian analysis, that of determining $f_R(\lambda)$, is equivalent to determining the values of α and β. There are a number of possibilities.

The most commonly used estimates are the method of moments estimates (MME), where α and β are chosen so that the mean and variance associated with the gamma distribution are equal to the mean and variance of the sample. That is, let $\bar{x}$ be the sample mean of the observed accident rates, s^2 be the sample variance of the observed accident rates, and m be the number of sites in the region. Then

$$\bar{x} = \frac{1}{m} \sum_{i=1}^{m} \frac{N_i}{V_i}$$
$$s^2 = \frac{1}{m-1} \sum_{i=1}^{m} \left(\frac{N_i}{V_i} - \bar{x}\right)^2$$

Using the MME, one selects α and β so that $\bar{x} = \alpha/\beta$ and $s^2 = \alpha/\beta^2$, or equivalently, $\beta = \bar{x}/s^2$ and $\alpha = \beta\bar{x}$.

Other commonly used estimates are the maximum likelihood estimates (MLE), where α and β are chosen so that they represent the values that are most likely to have generated the observed data. That is, if $\hat{\lambda}_i$ is the observed accident rate at site i (i.e., $\hat{\lambda}_i = N_i/V_i$), then α and β are chosen to maximize

$$\mathcal{L}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m \mid \alpha, \beta) = \prod_{i=1}^{m} \frac{\beta^\alpha}{\Gamma(\alpha)} \hat{\lambda}_i^{\alpha-1} e^{-\beta\hat{\lambda}_i}$$
$$= \left\{\frac{\beta^\alpha}{\Gamma(\alpha)}\right\}^m \left[\prod_{i=1}^{m} \hat{\lambda}_i\right]^{\alpha-1} e^{-\beta\sum_{i=1}^{m}\hat{\lambda}_i}$$

The function $\mathcal{L}$ represents the likelihood function associated with the observed data when the parameters α and β are assumed. The MLE values for α and β may be obtained by solving the equations

$$\frac{\partial \mathscr{L}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m \mid \alpha, \beta)}{\partial \alpha} = 0$$

$$\frac{\partial \mathscr{L}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m \mid \alpha, \beta)}{\partial \beta} = 0$$

Although the MME and the MLE are among the most commonly used methods of parameter estimation, other methods for estimating α and β exist and are discussed at length by Berger (*9*).

Once values for α and β have been determined, the first step of the analysis has been completed. In the second step, the observed accident rate at each site is used in combination with the gross estimate of the regional probability distribution to obtain the site-specific probability density functions, $f_i(\lambda \mid N_i, V_i)$. These density functions are obtained using Bayes's theorem. That is,

$$f_i(\lambda \mid N_i, V_i) \propto f(N_i \mid \lambda, V_i) f_R(\lambda)$$

Within the framework of Bayesian analysis, it is well known that under Assumptions A1 and A2, the resulting probability distribution $f_i(\lambda \mid N_i, V_i)$ is a gamma distribution (*9, 10*). Moreover, the parameters associated with this distribution, α_i and β_i, are easily obtained from the original choices of α and β and the observed data, N_i and V_i, as follows:

$$\alpha_i = \alpha + N_i$$

$$\beta_i = \beta + V_i$$

Thus, the probability density function associated with the accident rate at location $i(\tilde{\lambda}_i)$ is given by

$$f_i(\lambda \mid N_i, V_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i \lambda}$$

Note that as N_i and V_i increase, the site-specific parameters (α_i and β_i) will be largely determined by the observed data (N_i and V_i) and will become insensitive to the initial choice of α and β. As such, for each computation, it may be preferable to use the MME values rather than the MLE values, because they are substantially easier to calculate. All computations within this paper were based on the MME values of α and β.

With this collection of probability density functions, the identification of hazardous locations is now a straightforward matter. If $\bar{\lambda}$ is an upper limit on the "acceptable" accident rates, then we wish to identify a site i as hazardous if the probability is significant that $\tilde{\lambda}_i$ exceeds $\bar{\lambda}$. That is, if

$$P(\tilde{\lambda}_i > \bar{\lambda} \mid N_i, V_i) > \delta$$

where δ is some predetermined tolerance level, then site i is recognized as a hazardous location. Naturally, the appropriate values for $\bar{\lambda}$ and δ must be determined. For example, in the results section of this paper, various values of $\bar{\lambda}$ and δ are used to develop criteria for the identification of hazardous locations that are analogous to the criteria used in classically based statistical procedures. This allows a direct comparison between the results obtained from the Bayesian procedure presented in this paper and the results obtained from the classical techniques.

DATA DESCRIPTION

For the purposes of this study, 5-year (July 1981–June 1986) accident histories for signalized intersections under the jurisdiction of the Pima County Department of Transportation, in Tucson, Arizona, were used. Because significant improvement plans were undertaken during the third year (July 1983–June 1984), the data were broken into two separate sets. The first set corresponds to July 1981–June 1983, whereas the second set corresponds to July 1984–June 1986. Between July 1981 and June 1984, four intersections were signalized. Thus, the first data set includes 33 intersections, and the second data set includes 37 intersections. The two data sets were analyzed independently. For each intersection, the observed accident rate was calculated as the ratio of the total number of accidents to the total traffic volume over the 2-year period. The data are summarized in Tables 1 and 2.

One should note that the observed accident rate over each 2-year period is calculated as $N \times 10^6/2V \times 365$, and thus is normalized to represent the accident rate per million vehicles entering the intersection. The last two columns represent the probability that the site is hazardous on the basis of the two criteria developed in the section on results; these elements are discussed in further detail in that section.

EMPIRICAL RESULTS

In order to compare the results of an analysis based on classical statistical methods (e.g., those based on statistical confidence intervals) with the results of an analysis based on the Bayesian methodology, the two analyses must use analogous criteria in identifying a hazardous location. In practice, two commonly used criteria can be stated as follows:

C1. Site i is hazardous if the observed rate, $\hat{\lambda}_i$, exceeds the observed average rate across the region, x, with a level of confidence equal to δ.

C2. Site i is hazardous if the observed accident rate, $\hat{\lambda}_i$, exceeds the site's critical rate, which is a function of the observed regional accident rate, the traffic volume at site i, and the level of confidence desired, δ.

Typically, δ is a reasonably high number, such as 0.99, 0.95, or 0.90. C1 is the standard confidence-based crite-

TABLE 1 DATA OBTAINED FROM JULY 1981 THROUGH JUNE 1983

site number	observed accident rate (#/MVE)	number of accidents (N)	daily volume (V)	prob. (B1)	prob. (B2)
1	0.957	20	28644	0.4308	0.3861
2	1.192	46	52891	0.8684	0.8331
3	0.947	20	28950	0.4145	0.3701
4	1.437	43	40994	0.9813	0.9738
5	0.588	9	20965	0.0742	0.0614
6	1.043	14	18393	0.5402	0.5005
7	1.418	17	16422	0.8609	0.8377
8	0.779	9	15825	0.2573	0.2285
9	1.375	14	13953	0.8046	0.7776
10	1.007	18	24496	0.5058	0.4621
11	1.074	14	17863	0.5740	0.5349
12	1.174	21	24515	0.7280	0.6897
13	0.660	14	29054	0.0754	0.0608
14	1.040	22	28998	0.5632	0.5167
15	1.133	18	21773	0.6634	0.6237
16	0.675	11	22343	0.1198	0.1007
17	0.846	5	8100	0.3700	0.3411
18	0.742	11	20323	0.1897	0.1640
19	0.617	9	19989	0.0965	0.0810
20	0.282	4	19450	0.0061	0.0047
21	0.709	11	21253	0.1545	0.1318
22	1.003	8	10924	0.4805	0.4468
23	1.010	15	20360	0.5034	0.4626
24	0.088	1	15550	0.0025	0.0019
25	1.848	17	12605	0.9627	0.9543
26	0.567	6	14500	0.1138	0.0978
27	1.337	15	15376	0.7967	0.7683
28	1.471	46	42850	0.9891	0.9842
29	1.604	14	11957	0.8908	0.8727
30	1.032	15	19915	0.5311	0.4904
31	0.963	13	18502	0.4441	0.4054
32	1.184	20	23147	0.7308	0.6935
33	0.589	9	20953	0.0745	0.0616

rion, whereas C2 corresponds to the rate-quality criterion, developed by Norden et al. (*2*).

To identify hazardous locations by using Criterion C1, one must calculate both the sample mean $\bar{x}$ and the sample standard deviation s. Associated with each value of δ is a constant k_δ (e.g., $k_{0.95} = 1.645$), and if

$$\hat{\lambda}_i > \bar{x} + k_\delta s$$

then site i is said to be hazardous at the δ confidence level (*11*). To identify hazardous locations using the Bayesian methodology, a criterion that is analogous to C1 can be stated as follows:

B1. Site i is hazardous if the probability is greater than δ that its true accident rate, $\hat{\lambda}_i$, exceeds the observed average rate across the region.

Recall that the Bayesian methodology treats the accident rate at a particular location as a random variable and obtains a refined estimate of its probability distribution. As such, if

$$P\{\tilde{\lambda}_i > \bar{x} \mid N_i, V_i\} > \delta$$

then site i is said to be hazardous. Thus, the identification of hazardous locations using Criterion B1 involves the computation of

$$\begin{aligned} P\{\tilde{\lambda}_i > \bar{x} \mid N_i, V_i\} &= 1 - P\{\tilde{\lambda}_i \leq \bar{x}\} \\ &= 1 - \int_0^{\bar{x}} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i \lambda} d\lambda \qquad (1) \end{aligned}$$

If the computed value exceeds δ, site i is identified as hazardous.

Similarly, to identify hazardous sites using Criterion C2, one must calculate the regional accident rate,

$$x_R = \frac{\sum_i N_i}{\sum_i V_i}$$

For a given level of confidence, δ, the critical rate associated with location i is computed as follows:

$$\lambda c_i = x_R + k_\delta \sqrt{\frac{x_R}{V_i}} + \frac{1}{2V_i}$$

TABLE 2 DATA OBTAINED FROM JULY 1984 THROUGH JUNE 1986

site number	observed daily rate (#/MVE)	number of accidents (N)	daily volume (V)	prob. (B1)	prob. (B2)
1	0.710	15	28950	0.0727	0.0616
2	1.006	51	69450	0.3890	0.3405
3	1.047	27	35350	0.4845	0.4458
4	0.897	22	33600	0.2398	0.2118
5	1.051	23	30000	0.4885	0.4522
6	0.625	12	26300	0.0414	0.0347
7	0.875	16	25050	0.2443	0.2190
8	0.726	11	20750	0.1266	0.1114
9	0.865	13	20600	0.2544	0.2303
10	0.648	13	27500	0.0463	0.0388
11	1.028	18	24000	0.4520	0.4191
12	1.357	36	36350	0.9083	0.8903
13	1.237	31	34350	0.7828	0.7528
14	1.162	24	28300	0.6560	0.6222
15	0.704	15	29200	0.0681	0.0575
16	0.689	14	27850	0.0648	0.0548
17	0.857	6	9592	0.3238	0.3033
18	1.109	27	33350	0.5909	0.5533
19	1.318	28	29100	0.8436	0.8202
20	0.386	8	28400	0.0017	0.0013
21	1.342	29	29600	0.8669	0.8456
22	1.269	15	16200	0.7032	0.6770
23	0.947	18	26050	0.3340	0.3036
24	0.684	12	24050	0.0790	0.0679
25	2.289	34	20350	0.9998	0.9997
26	1.006	13	17700	0.4275	0.3984
27	1.300	24	25300	0.8034	0.7782
28	2.177	66	41550	1.0000	1.0000
29	1.912	30	21500	0.9951	0.9937
30	1.288	22	23400	0.7795	0.7534
31	0.988	16	22200	0.3983	0.3675
32	1.266	20	21650	0.7445	0.7173
33	1.160	26	30700	0.6630	0.6284
34	0.601	5	11400	0.1361	0.1230
35	0.813	11	18550	0.2145	0.1935
36	0.525	12	31300	0.0081	0.0064
37	0.602	9	20500	0.0569	0.0489

This critical rate is based on the assumption that the number of accidents at location i is Poisson distributed with a mean of $x_R V_i$ (*2*), which is similar to Assumption A1 in this paper. The critical rate is defined so that with this assumption, the observed accident rate will be less than or equal to the critical rate with probability δ. An investigation into the early development of this rate-quality method (*2*) suggests that an analogous criterion within the Bayesian methodology can be stated as follows:

B2. Site i is hazardous if the probability is greater than λ that its accident rate, $\tilde{\lambda}_i$, exceeds the observed regional accident rate, x_R.

That is, under Criterion B2, site i is identified as hazardous if

$$P\{\tilde{\lambda}_i > x_R \mid N_i, V_i\} > \delta$$

or equivalently, if

$$1 - \int_0^{x_R} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i \lambda} d\lambda > \delta \qquad (2)$$

Note that the assumptions leading to Criteria C2 and B2, namely, those regarding the Poisson nature of accidents at a particular site, are very similar. The fundamental difference lies in the fact that in using C2, one implicitly assumes that the true accident rate is x_R (*2*). The authors who pioneered this method concede that the true rate (*2*) "is never known and we shall always have to be satisfied with an estimate of the expectation" (i.e., x_R). In using Criterion B2, one accounts for the inherent randomness associated with each accident rate, as reflected in Assumption A2.

In identifying hazardous locations on the basis of the Bayesian methodology (i.e., Criteria B1 and B2), one must perform the integrations identified in Equations 1 and 2. A computer program was written to numerically evaluate each of these integrals. The results of our empirical study are summarized in Figures 1 through 4. For each data set and for each value of δ (i.e., $\delta = 0.99$, $\delta = 0.95$, and $\delta = 0.90$), hazardous sites were identified on the basis of Criteria C1 and C2, corresponding to the classical statistical methods, and of the analogous Bayesian Criteria B1 and B2. These results are presented in Figures 1–4. The ele-

a. $\delta = 0.99$

	HC1	NC1
HB1	0	0
NB1	0	33

b. $\delta = 0.95$

	HC1	NC1
HB1	1 (25)	2 (4,28)
NB1	1 (29)	29

c. $\delta = 0.90$

	HC1	NC1
HB1	2 (25,28)	1 (4)
NB1	1 (29)	29

FIGURE 1 Distribution of sites based on B1 and C1 (July 1981–June 1983).

a. $\delta = 0.99$

	HC2	NC2
HB2	0	0
NB2	2 (25,28)	31

b. $\delta = 0.95$

	HC2	NC2
HB2	3 (4,25,28)	0
NB2	0	30

c. $\delta = 0.90$

	HC2	NC2
HB2	3 (4,25,28)	0
NB2	2 (7,29)	28

FIGURE 2 Distribution of sites based on B2 and C2 (July 1981–June 1983).

a. $\delta = 0.99$

	HC1	NC1
HB1	2 (25,28)	1 (29)
NB1	0	34

b. $\delta = 0.95$

	HC1	NC1
HB1	3 (25,28,29)	0
NB1	0	34

c. $\delta = 0.90$

	HC1	NC1
HB1	3 (25,28,29)	1 (12)
NB1	0	33

FIGURE 3 Distribution of sites based on B1 and C1 (July 1984–June 1986).

a. $\delta = 0.99$

	HC2	NC2
HB2	3 (25,28,29)	0
NB2	0	34

b. $\delta = 0.95$

	HC2	NC2
HB2	3 (25,28,29)	0
NB2	0	34

c. $\delta = 0.90$

	HC2	NC2
HB2	3 (25,28,29)	0
NB2	1 (12)	33

FIGURE 4 Distribution of sites based on B2 and C2 (July 1984–June 1986).

ments in the 4×4 matrices found in Figure 1 are organized as follows:

1. Columns

a. HC1 corresponds to the number of sites that were identified as hazardous on the basis of Criterion C1 (i.e., the number of sites whose observed accident rate exceeded $\bar{x} + k_\delta s$).

b. NC1 corresponds to the number of sites that were not identified as hazardous on the basis of Criterion C1.

2. Rows

a. HB1 corresponds to the number of sites that were identified as hazardous on the basis of Criterion B1 (i.e., the number of sites with $P\{\tilde{\lambda}_i > \bar{x}\} \geq \delta$).

b. NB1 corresponds to the number of sites that were not identified as hazardous on the basis of Criterion B1.

The numbers in parentheses correspond to the sites that are identified as hazardous, as represented in Tables 1 and 2.

Thus, for example, in Figure 1 we see that for July 1981–June 1983, for $\delta = 0.95$, one site (25) is identified as hazardous under both C1 and B1, two sites (4, 28) are identified as hazardous under B1 but not C1, and one site (29) is identified as hazardous under C1 but not B1. The remaining 29 sites are not identified as hazardous under either criterion.

The results of the comparison between C2 and B2 for July 1981–June 1983 are similarly arranged in Figure 2 and those for the analyses of the data collected between July 1984 and June 1983 are presented in Figures 3 and 4 for B1 and C1 and for B2 and C2, respectively.

Because there seems to be consistent disagreement between the various criteria, a discussion of the information conveyed in Figures 1–4 is in order. First, note that as expected, as δ decreases, the number of sites identified increases under all four criteria. That is, the more relaxed the identification requirement, the easier it is to be identified.

Second, there is very little difference between the sites identified by B1 and B2, the Bayesian criteria. The only difference is in Site 12, using the data collected from July 1984–June 1986. With $\delta = 0.90$, it is identified using B1 but not B2 (see Figures 3 and 4). However, the data presented in Table 2 indicate that the probability computed under Criterion B1 is 0.9083, whereas the probability computed under Criterion B2 is 0.8901. Thus, although the methods differ, the difference is not substantial. This significant agreement between the two methods is easily explained by the data. B1 uses the threshold value $\overline{\lambda} = \bar{x}$, whereas B2 uses $\overline{\lambda} = x_R$. For both data sets, x and x_R are not substantially different, as indicated by the summary statistics presented in Table 3.

One should note carefully that Criteria B1, B2, and C2 all tend to be more conservative than C1, in that they tend to identify more sites as hazardous. This suggests that C1 may be more susceptible to the identification of false negatives (i.e., those sites that are actually hazardous but are not identified as such).

In a review of Figures 3 and 4, it is clear that for the data collected between July 1984 and June 1986, the classical criteria are in relatively high agreement with their Bayesian counterparts. This is most likely due to the extremely high accident rates (2.289, 2.177, and 1.912, respectively) of three intersections (25, 28, and 29), compared with a mean accident rate of 1.0396 and a regional rate of 1.0578. Because of the extreme nature of these three accident rates, it is reasonable to expect that any justifiable procedure would identify these sites as hazardous and that all others may seem safe by comparison. Of course, when the Bayesian procedure is used, a change in the threshold value would affect the sites that are identified as hazardous. For the purposes of this study, the values of $\overline{\lambda} - \bar{x}$(B1) and $\overline{\lambda} - x_R$(B2) were chosen so that the Bayesian and classical methods could be easily compared.

TABLE 3 SUMMARY STATISTICS

Data Set	$\bar{x}$	x_R	s
June 1981 - July 1983	0.9815	1.0042	0.3756
June 1984 - July 1986	1.0396	1.0578	0.4196

The disagreement between Criteria B1 and C1 (Figures 1 and 3) is probably best explained in terms of the underlying assumptions. B1 is based on the widely accepted assumptions that accidents occur according to a Poisson distribution and that the accident rate has a gamma distribution (i.e., Assumptions A1 and A2 as stated in the section on Bayesian methodology in this paper). Similarly, Criterion C1 is based on the implicit assumption that the observed accident rates are normally distributed. The relatively large standard deviations when compared with the low sample means (e.g., $\bar{x} = 0.9815$ and $s = 0.3756$ for the July 1981–June 1983 data) combined with the fact that the accident rates must be nonnegative suggest that the normal distribution may yield an inappropriate model for these data sets.

Because Criteria B2 and C2 are based on a similar assumption (i.e., that accidents occur according to a Poisson distribution), the differences between the corresponding results, as summarized in Figures 2 and 4, are due solely to the treatments of the actual accident rate. The Bayesian method explicitly assumes that the accident rate at any given site is a random variable and accounts for this randomness in the identification process. The rate-quality method (C2) implicitly assumes that the accident rate at each site is equal to the regional rate x_R. Thus, the Bayesian method allows site-specific accident information to guide the identification process, and the rate-quality method does not.

Finally, it should be noted that Intersection 4, which was identified as hazardous by using the data collected from July 1981–June 1983 with $\delta = 0.95$ under Criteria B1, B2, and C2, underwent significant change during July 1983–June 1984. Subsequently, it was no longer identified as a hazardous intersection, and the probability that it is hazardous dropped from 0.9813 to 0.2398 on the basis of B1 or from 0.9738 to 0.2190 on the basis of B2. Clearly, these substantial drops indicate that the improvement program was successful.

CONCLUSIONS

Use of a Bayesian analysis in the identification of hazardous accident locations using accident rate data appears to be a fundamentally sound procedure, which is shown to have identification criteria analogous to those used in the classical identification scheme, although it is certainly not limited to these criteria. The Bayesian technique has the added advantage of allowing the assessment of the impact of varying the degree of confidence, δ, without requiring that the decision statistics be recomputed. Moreover, knowing the probability that the actual rate exceeds the regional rate, for example, provides added information that can be used to evaluate the trade-offs involved in deciding which sites are candidates for improvement funds.

The results presented in Figures 1 and 3 suggest that, in general, the confidence-based procedure, C1, may be inappropriate for identifying hazardous locations. Criterion C1 fails to identify as hazardous many sites that are flagged by Criteria B1, B2, and C2. The underlying assumption of normality in the distribution of the accident rate appears to cause C1 to err in the direction of false negatives. This is the least desirable characteristic for an identification procedure.

The results presented in Figures 2 and 4, combined with an analysis of the underlying assumptions, suggest that in many cases, the use of B2 may be preferable to the use of C2. This may be especially true when data are sparse or when numerous years of comparable data are not available. It is expected that the differences between B2 and C2 will be substantially reduced whenever a great deal of data are available.

REFERENCES

1. D. A. Morin. Application of Statistical Concepts to Accident Data. In *Highway Research Record 187,* HRB, National Research Council, Washington, D.C., 1967, pp. 72–79.
2. M. Norden, J. Orlansky, and H. Jacobs. Application of Statistical Quality-Control Techniques to Analysis of Highway-Accident Data. In *Bulletin 117*, HRB, National Research Council, Washington, D.C., 1956, pp. 17–31.
3. C. P. Brinkman. Toward Improved Highway Safety Analysis. *Public Roads,* Vol. 50, No. 2, 1986, pp. 49–54.
4. E. Hauer. On the Estimation of the Expected Number of Accidents. *Accident Analysis and Prevention,* Vol. 18, No. 1, 1986, pp. 1–12.
5. E. Hauer and B. Persaud. Common Bias in Before-and-After Accident Comparisons and Its Elimination. In *Transportation Research Record 905,* TRB, National Research Council, Washington, D.C., 1983, pp. 164–174.
6. E. Hauer and B. Persaud. The Problem of Identifying Hazardous Locations Using Accident Data. In *Transportation Research Record 905,* TRB, National Research Council, Washington, D.C., 1984, pp. 36–43.
7. B. Persaud and E. Hauer. A Comparison of Two Methods for De-Biasing Before-and-After Accidents Studies. In *Transportation Research Record 975,* TRB, National Research Council, Washington, D.C., 1984, pp. 43–49.
8. W. D. Glauz, K. M. Bauer, and D. J. Migletz. Expected Traffic Conflict Rates and Their Use in Predicting Accidents. In *Transportation Research Record 1026,* TRB, National Research Council, Washington, D.C., 1985, pp. 1–12.
9. J. O. Berger. *Statistical Decision Theory and Bayesian Analysis,* 2nd ed. Springer Verlag, New York, 1985.
10. M. D. DeGroot. *Optimal Statistical Decisions.* McGraw-Hill, New York, 1970.
11. J. C. Laughlin, L. E. Haefner, J. W. Hall, and D. R. Clough. *NCHRP Report 162: Methods for Evaluating Highway Safety Improvements.* TRB, National Research Council, Washington, D.C., 1975.

DISCUSSION

CARL N. MORRIS

Department of Mathematics and Statistical Sciences Center, University of Texas, Austin, Tex. 78712

The proposal by Higle and Witkowski to use hierarchical Bayesian and empirical Bayesian methods to identify hazardous highway locations is a sound one that can be expected to fare well in comparison with more traditional methods. Accident rate estimation is extremely uncertain because the number of accidents at any one intersection tends to be quite random and subject to the regression-to-the-mean phenomenon. Bayesian methods help because they effectively permit pooling of data from other relevant sites. It also makes good sense that the authors base hazard determinations on probabilistic assessments of the value $\tilde{\lambda}_i$ of the intrinsic hazard rate, as, for example, Criteria B1 and B2 require.

In Table 1 the exposure rates (daily volumes) V_i vary by a factor of 6.5 between the two extreme sites (Sites 2 and 17). Although the setting here is for Poisson data, this is otherwise analogous to the empirical Bayes estimation of means in the normal distribution case, which is now well understood (*1, 2*). In such cases the Bayesian ranking of extreme intersections differs from the observed daily rates because those extreme intersections with low volumes generally would regress to the mean more than high-volume intersections. Thus, Site 4 in Table 1 probably is more hazardous than Site 25, even though the observed rate of 1.437 MVE is less than 1.848 MVE for Site 25. This occurs because the daily volume for Site 4 is more than three times higher, 40,994 to 12,605.

It should be obvious that a methodology that properly weighs all evidence in ranking dangerous intersections is very valuable. Such features can only be revealed by using Bayesian and empirical Bayesian models, in which distributions are specified and estimated for both observed data and unobserved parameters.

Despite the virtues of the authors' general idea, there are features in their proposed methods that need further adjustment to correct for bias and to improve statistical efficiency. I will explain the difficulties partly on the basis of my own research with Olga Pendleton on accident analysis using these same Poisson-gamma models.

The authors' development fits within the General Model for Statistics (*2, 3*), which in their Poisson-gamma setting and notation is summarized in Table 4.

Note that Equations 3 and 4 correspond to the authors' Assumptions A1 and A2. The descriptive model is entirely equivalent to the inferential model (Table 5), which reverses the probabilistic conditioning and is more convenient for statistical analysis.

Expressions 5 and 7 refer to the negative binomial and gamma distributions, with their usual parameterization, whereas Expressions 6 and 8 and the square brackets signify that the means and variances are displayed.

TABLE 4 DESCRIPTIVE MODEL

Observed Data:	$N_i \mid \tilde{\lambda}_i \sim \text{Poisson}(V_i\tilde{\lambda}_i)$ $i = 1, \ldots, m$ independently.	(3)
Unobserved Parameters:	$\tilde{\lambda}_i \mid (\alpha, \beta) \sim \text{Gamma}\left(\alpha, \frac{1}{\beta}\right)$ $i = 1, \ldots, m$ independently $\phi \equiv (\alpha, \beta)$ unknown, $\alpha, \beta > 0.$	(4)

TABLE 5 INFERENTIAL MODEL

Observed Data:	$N_i \mid \alpha, \beta \sim \text{NegBin}\left(\alpha, p_i \equiv \dfrac{V_i}{V_i + \beta}\right)$	(5)
	$= \text{NegBin}\left[\dfrac{\alpha V_i}{\beta}, \dfrac{\alpha V_i}{\beta} + \dfrac{\alpha V_i^2}{\beta^2}\right]$	(6)
	$i = 1, \ldots, m$ independently.	
Unobserved Parameters:	$\tilde{\lambda}_i \mid N_i, \alpha, \beta \sim \text{Gamma}\left(\alpha + N_i, \dfrac{1}{\beta + V_i}\right)$	(7)
	$= \text{Gamma}\left[\dfrac{\alpha + N_i}{\beta + V_i}, \dfrac{\alpha + N_i}{(\beta + V_i)^2}\right]$	(8)
	$i = 1, \ldots, m$ independently.	

The analysis proceeds using $(N_i, \ldots, N_m)$ to estimate $\phi = (\alpha, \beta)$ from Expression 5, and then carries this information to 7 or 8 to assess the posterior distribution. The simplest method for doing this, often called "empirical Bayes," simply develops a point estimate $\hat{\phi} = (\hat{\alpha}, \hat{\beta})$ and substitutes these values into 7 or 8. [This can be risky if $(\hat{\alpha}, \hat{\beta})$ are not accurately estimated, an issue that could be assessed in the data example.] The authors follow this empirical Bayes approach, although not quite correctly.

Note that the marginal distribution (Expression 5) for the data N_i is negative binomial, not gamma, as the authors indicate when discussing the MLE. Thus, the maximum likelihood conditions in the second section of their paper are incorrect.

Similarly, the MME technique is improperly applied by the authors, and the estimates of the "hyperparameter" $\phi = (\alpha, \beta)$ are biased and inefficient. Define $X_i = N_i/V_i$, $X = 1/m\Sigma X_i$ and $s^2 = [1/(m - 1)]\Sigma(X_i - \bar{X})^2$, as in the second section. Then from Expression 6, $EX_i = \alpha/\beta$ and so $E\bar{X} = \alpha/\beta$, as claimed. However, from Expression 6,

$$\text{Var}(X_i) = \frac{\alpha}{\beta V_i} + \frac{\alpha}{\beta^2} \equiv \sigma^2 \tag{9}$$

and so

$$Es^2 = \frac{1}{m} \sum_1^m \sigma_i^2 \tag{10}$$

$$= \frac{\alpha}{\beta^2} + \frac{\alpha}{\beta V^*} \tag{11}$$

This exceeds α/β^2, the value claimed by the authors, by the amount $\alpha/\beta V^*$, V^* the harmonic mean of $(V_i, \ldots, V_m)$. It follows that, instead of the authors' formula, the MME is

$$\hat{\beta} = \frac{V^* \bar{X}}{V^* s^2 - \bar{X}} \qquad \hat{\alpha} = \bar{X}\hat{\beta} \tag{12}$$

Of course additional modifications are required if the denominator of β is not positive or is close to zero.

For the data set of Table 1 (1981–1983), the authors' formulas give values for α and β) of about 7 and 7, respectively, but the correct MME equations give estimates closer to 14 and 14. Thus, the model in Expression 4 should provide about twice as much information as the authors have estimated.

We could further improve the estimation. The unweighted mean $\bar{X}$ is not the best estimate of α/β because those X_i based on larger exposures V_i deserve more weight. A similar statement would apply to the use of s^2, but determining the correct weights is quite complicated in that case. The benefits of these improvements with m = 33 or 37, as in the examples, should be of second order, however, and so perhaps this use of a simple method is not costly for the data considered.

A nice feature of using the optimally weighted $\bar{X}$ is that it then could be used instead of $\bar{X}$ and X_R in Criteria B1 and B2 and would be preferable to either.

The simple device of substituting the estimates (α, β) from Expression 5 into 7 fails to acknowledge the uncertainty in knowledge of (α, β). More accurate methods, which are hard to derive [see discussion by Morris (*3*)] would spread out the posterior distribution. The effect of properly accounting for this in the Higle-Witkowski application would be to lower somewhat the probabilities for Criteria B1 and B2 for those locations with high accident rates.

To summarize, the Higle-Witkowski Bayesian model promises to have many advantages over standard methods for identifying hazardous locations. It will take more time before the most appropriate analytical methods are available, however. The development of such methods promises to be a rewarding and interesting task.

ACKNOWLEDGMENT

This discussion was based on research supported by FHWA at the Texas Transportation Institute, Texas A&M University, College Station, Texas. Olga Pendleton is the principal investigator.

REFERENCES

1. B. Efron and C. N. Morris. Data Analysis Using Stein's Estimator and Its Generalizations. *Journal of the American Statistical Association,* Vol. 70, 1975, pp. 311–319.
2. C. N. Morris. Parametric Empirical Bayes Inference: Theory and Applications (with discussion). *Journal of the American Statistical Association,* Vol. 78, 1983, pp. 47–65.
3. C. N. Morris. Determining the Accuracy of Bayesian Empirical Bayes Estimators in the Familiar Exponential Families. In *Statistical Decision Theory and Related Topics, IV,* Vol. 1 (S. S. Gupta and J. O. Berger, eds.), Springer-Verlag, New York, 1988, pp. 251–263.

The opinions expressed in this discussion are those of the author alone, and have not been reviewed by FHWA.

DISCUSSION

OLGA J. PENDLETON
Texas Transportation Institute, College Station, Tex. 77843.

Morris, in his discussion of this paper, has eloquently and didactically covered its weaknesses. The following discussion is offered to paraphrase and reemphasize some of his comments and to add a few comments regarding the numerical example.

Basically, this study uses a method that is inefficient. Furthermore, there are computational errors in the analysis. The authors' estimates are biased and inefficient and the method of moments technique is improperly applied, as Morris has noted. With regard to the computational errors, Morris has shown that the discrepancy in the parameter estimates is twofold when the computations are done correctly.

My second comment pertains to the numerical example. Although the results are computationally incorrect, I will base my comments on the data as they were originally presented, to show that many of the authors' claims are not supported by their numerical example.

The authors attempt to show by the numerical example that the empirical Bayesian (EB) methods (Criteria B1 and B2) are superior to two classical methods—one that assumes a normal distribution (C1) and one that assumes the more correct Poisson distribution (C2). Careful inspection of the results (Figures 1–4) does not support this claim.

In Figure 1 the classical estimate, C1, identifies one site as hazardous that the EB method fails to recognize (Site 29), and fails to identify one site that the EB method does identify (Site 4). Note that the accident rate for the site identified by the classical method is higher (1.604) than that identified by the EB method (1.437), and hence might appear to be "more logical."

In Figure 2 the classical estimate using the more correct Poisson distribution assumption, C2, identifies the same hazardous sites as does the comparable EB method (B2) and recognizes them at an even higher δ than does the EB method. The classical method identifies two sites (7 and 29) that EB does not at $\delta = 9$. Thus, one could rephrase the authors' first sentence in the second paragraph of page 5 in support of the classical estimator as follows: "Criterion B2 (EB) fails to identify as hazardous many sites that are flagged by C1 and C2." The "many" here would be only three sites; however, this is the same number of sites the authors refer to as "many" in their original statement denouncing C1 (Figure 1).

In Figures 3 and 4 we see one site (Figure 3) that the EB method identified and the classical one did not, namely, Site 12, but then in Figure 4, the classical method identifies Site 12 as hazardous when the EB method does not.

In summation, this numerical example does not show, as stated in the abstract, "that some classically based statistical techniques may be prone to err in the direction of false negatives" any more than EB methods. If, after correcting the computational error discovered by Morris, this example still fails to support such a claim, a better example should be found. Otherwise, the would-be user of this methodology is left to conclude that here we have a much more complicated procedure that does no better (maybe even worse) than the more simplistic classical methods. In its present form, this paper appears to do a disservice to a methodology that may, in fact, be superior by (a) containing mathematical errors and (b) presenting a weak example.

DISCUSSION

EMMANUEL QUAYE
Department of Civil Engineering, University of Toronto, Box 10, 321 Bloor Street West, Toronto, Ontario M5S 1S5, Canada.

In this paper the authors suggest ways of estimating the mean and the variance of true accident rates for m sites. The expressions given in the paper for calculating the variance by the method of moments and the formulation of the likelihood function are incorrect.

It can be shown (*1*) that, using the method of moments,

$$\hat{\mathrm{Var}}\{\lambda\} = [1/(m-1)]\left[\sum_1^m (N_i^2 - N_i)/V_i^2 - (1/m)\left(\sum_1^m N_i/V_i\right)^2\right] \quad (13)$$

and not

$$[1/(m-1)] \sum_1^m \left(N_i/V_i - 1/m \sum_1^m N_i/V_i\right)^2 \quad (14)$$

as given in the paper.

As can be seen from the expressions above, Equation 2 leads to an overestimation of the variance. The difference between the two expressions is

$$[1/(m-1)] \sum_1^m (N_i/V_i^2) \quad (15)$$

Using data from Tables 1 and 2, the following comparisons, as shown below, can be made. It can be observed that there is a substantial overestimation of the variance by Equation 2.

	Variance	
Data Set	*Equation 1*	*Equation 2*
Table 1	0.0673	0.1410
Table 2	0.1164	0.1759

Similarly, it can be shown (*1*) that the correct likelihood function is

$$\prod_1^m [\alpha/(\alpha + V_i E\{\lambda\})]^{\alpha}[\Gamma(\alpha + N_i)/(\Gamma(\alpha)N_i!)] \times [V_i E\{\lambda\}/(\alpha + V_i E\{\lambda\})]^{N_i} \quad (16)$$

and not the corresponding expression given in the paper. The performance of this likelihood function was confirmed by simulation (*1*) and shown to give good results.

REFERENCE

1. E. Hauer and P. Garder. Research Into the Validity of the Traffic Conflict Technique. *Accident Analysis and Prevention*, Vol. 18, 1986, pp. 471–481.

AUTHORS' CLOSURE

First, let us state unequivocably that it is truly an honor to have these authors share their thoughts on this paper. It is with pleasure that we respond to their comments.

In our paper, we present a two-stage method for obtaining information about accident rates in a Bayesian fashion. In the first stage, we combine accident histories from various sites within a region to estimate a regional distribution, whereas in the second stage, we use site-specific data to update the regional distribution, thereby obtaining refined estimates of the distribution associated with each site. The comments in the discussions offered by Quaye and Morris concern some of the details associated with the first stage. We shall address their comments first, and save our discussion of Pendleton's criticisms for last.

FIRST-STAGE CONCERNS

Both Quaye and Morris question the manner in which we compute our estimate of the variance of the regional distribution, $f_R(\lambda)$, which has an obvious impact on our initial choice for the parameters α and β. Quaye's objection arises from the fact that we have treated the observed accident rates $\hat{\lambda}_i = N_i/V_i$, $i = 1, \ldots, m$, as our sample of observations instead of the collection of paired values (N_i, V_i). Morris points out that even when the observed rates are used, the manner in which we calculate the sample variance yields a biased estimate of the distributional variance.

Admittedly, because each observed accident rate, $\hat{\lambda}_i$, is derived from two pieces of data, N_i and V_i, Quaye's interpretation of the "sample" may be preferred to ours. However, the method of estimating the variance of the accident rates, presented by Hauer and Garder (*1*), should be considered with caution. These authors verify that the estimate presented by Quaye can provide negative estimates of the variance, a quantity that is necessarily nonnegative. This can cause difficulties, and thus the procedure should be carefully examined before it is used. Although the estimate we used cannot yield a negative sample variance, it does yield a biased estimate of the true variance, as discussed by Morris.

Regardless, herein lies a major difficulty associated with some classical identification techniques. Each of these three methods represents a reasonable or common method used to estimate the variance of the regional distribution, yet each provides a different estimate (although the Quaye and Morris estimates are in very close agreement). Because of the direct dependence of the classical techniques on the computed sample mean and variance, it is clear that the resulting set of sites identified as hazardous depends on the manner in which the data are presented and the statistics are computed. Different estimates of the variance of the accident rates across the region will necessarily lead to different sets of sites that are identified as hazardous. As a result, the tremendous differences in the tabulated values of the computed estimates of the variance presented in the Quaye discussion, corresponding to factors of 2.09 and 1.51 for the first and second data sets, respectively, might cause some concern for the integrity of some of the classical procedures.

To see that the Bayesian technique does not suffer from these same shortcomings, one need only compare the distributions obtained with the two estimation techniques. Clearly, different estimates of the variance of the accident rates will result in different estimates of both the regional and the refined distributions. Figures 5 and 6 show the regional distribution and a representative refined distribution that result from the two estimation techniques. These figures are based on the moments computed from the second data set, because it best shows the differences between the various distributions. The more peaked curve in Figure 5 (i.e., the one with the smaller variance) corresponds to the regional distribution based on the estimate of the variance obtained with Quaye's Equation 13, which is very nearly equal to Morris's estimate. In Figure 6, one can see that the refined distributions are virtually indistinguishable. As we discussed in the paper, this is because in

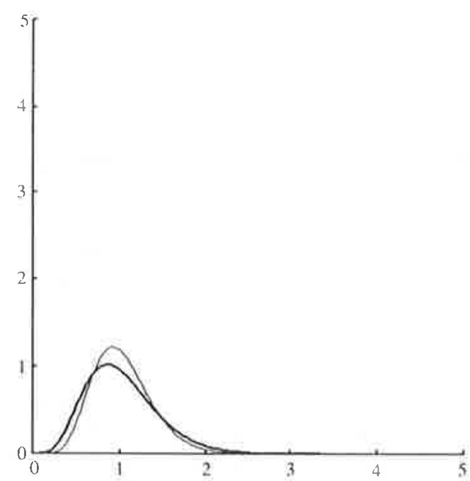

FIGURE 5 Regional distribution (July 1984–June 1986).

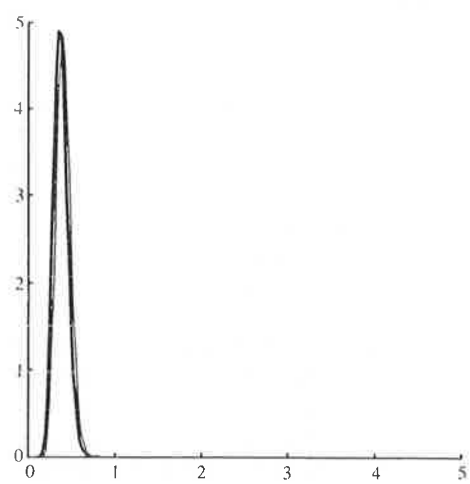

FIGURE 6 Refined distribution (July 1984–June 1986).

updating the parameters associated with the regional distribution to obtain the site-specific distribution, one necessarily overpowers the original parameters (which in our case are based on the computed sample mean and variance) with the site-specific data. Thus, the refined distributions, which provide the basis for the identifications and are therefore of paramount importance, are largely insensitive to the original parameter selection. It follows that one can be reasonably assured that although the classical methods are highly dependent on the sample variance, the Bayesian methods are not.

In addition, both Quaye and Morris question the likelihood function presented in the second section of the paper. In performing our analysis, we worked exclusively with the model that Morris has described as the descriptive model. As such, we have explicitly dealt only with the distribution of the accident rates, which are initially assumed to follow a gamma distribution. This is our Assumption A2. As a result, the likelihood function for the gamma distribution is correct as stated. Naturally, if we had performed our investigation on the basis of N_i, using the inferential model, the negative binomial model suggested by Morris would have been correct, and the likelihood function would change accordingly.

Unfortunately, Quaye's claim regarding the likelihood function (i.e., Quaye's Equation 16) is simply incorrect. A likelihood function is a mathematical entity representing the relative likeliness of the observed data (e.g., $\{(N_i, V_i)\}_{i=1}^{m}$ for a given set of distribution parameters (e.g., α and β for a gamma distribution). Thus, the maximum likelihood parameter estimates for the gamma distribution are those values of α and β that are most likely to have generated the observed data. In a very real sense, they provide the values of α and β that best fit the observed data, although the resulting theoretical mean and variance need not agree with the sample mean and variance. The "likelihood function" offered by Quaye, which also appears in the paper by Hauer and Garder (*1*), is not a true likelihood function. Instead, it is a form of a likelihood function that has been artificially constrained so that the resulting theoretical mean agrees with the observed sample mean. Thus, in general, the parameter estimates that are obtained by using it are not the maximum likelihood estimates, whereas those obtained from the expression in our paper are. In addition, the statement that somehow Quaye's Equation 16 has been "confirmed" by simulation is most disturbing indeed. As a mathematically known function, a likelihood function need not be subjected to empirical validation. Such a validation procedure suggests that there may be a "gray area" associated with the functional definition. Because it is a well-known mathematical entity, there is no such gray area requiring empirical validation.

INTERPRETATIONAL CONCERNS

Pendleton's discussion begins with a simple reiteration of Morris's comments. Because we have already discussed these comments at length, there is no need to further expand on them here. Instead, we shall focus on Pendleton's remaining criticism, which pertains to our interpretation of the results of our empirical study.

First, note that we are not in a position to claim that one technique is superior to another. To do so, one would have to know which sites are actually hazardous so that one can correctly determine which technique tends to provide correct identifications most often. Naturally, in the absence of perfect information, one can only interpret pieces of evidence or results as they become available.

The interpretation of empirical results such as those presented in our paper is necessarily subjective, and solid conclusions are often difficult to reach. Our conjecture that the classical technique, C1, "may be prone to err in the direction of false negatives" is based on an in-depth analysis of the magnitude of the differences in the levels at which sites are identified as hazardous.

To illustrate these differences, consider the three sites from Figure 1 for which techniques B1 and C1 provide differing results, namely, Sites 2, 28, and 29. From Table 1 one can obtain the probabilities computed using the Bayesian technique. A simple algebraic expression identifies the maximum confidence level at which Criterion C1 will identify these sites as hazardous, δ_{max}. Similar quantities can be obtained from the analogous sites associated with Figure 3 (i.e., Sites 12 and 29). These values are summarized in Table 6.

On the basis of this information, it seems clear that B1 provides very strong evidence that Sites 4 and 28 are hazardous (0.9813 and 0.9891, respectively), whereas C1 provides substantially weaker evidence (0.8874 and 0.9037, respectively). Similarly, although Site 12 receives a lower degree of support from B1 than do Sites 4 and 28, it still receives a substantially higher level of support from B1 (0.9083) than from C1 (0.7753). Of course, in the first data set, Site 29 receives a lower level of support from B1 (0.8908) than from C1 (0.9512), but the difference is smaller in this case. Because we believe that this type of analysis provides a better understanding of the difference between the methods than does Pendleton's method, we stand by our earlier claim. Of course, further investigation of the differences between the Bayesian and classical meth-

TABLE 6 SUMMARY OF DIFFERENCES BETWEEN B1 AND C1

Figure 1:	site	prob. (B1)	δ_{max} (C1)
	4	0.9813	0.8874
	28	0.9891	0.9037
	29	0.8908	0.9512
Figure 3:	site	prob. (B1)	δ_{max} (C1)
	12	0.9083	0.7753
	29	0.9951	0.9812

ods is called for. In addition, because Pendleton's observation regarding the apparent reordering of Sites 4 and 29 in Figure 1 is eloquently explained in Morris's discussion, we shall not endeavor to expand on his explanation here.

In conducting the research reported in this paper, it was our intention to offer a Bayesian technique for identifying hazardous intersections and to begin to understand how our technique differs from some of the classical techniques. It was not, as Pendleton states, to show that the Bayesian methods "are superior to two classical methods." Pendleton suggests that because our data set fails to provide evidence of superiority of the Bayesian method, another example should be "found" that will support such a claim. The purpose of this research was to explore the truths of a situation, not to discard or manufacture data in an effort to support a desired result.

CONCLUSION

In conclusion, we would like to agree with the closing remarks made by Morris. The application of Bayesian analyses to accident data does appear to provide a fruitful avenue of exploration. There are numerous modeling techniques to be explored. In addition, a further understanding of the differences in the results provided by Bayesian and classical identification techniques is of obvious importance. The pursuit of knowledge in this exciting field promises to offer its own rewards.

Do Traffic Signals Affect Safety? Some Methodological Issues

Bhagwant N. Persaud

This paper is based on a review of often-cited studies on the safety impact of traffic signal installation. Most of these studies are found wanting with respect to methods of analysis or inferences from the results. Two common pitfalls—regression to the mean and incorrect inferences from cross-section studies—are illustrated. The first might lead to erroneous conclusions about the circumstances under which signal installation is likely to improve or degrade safety. The second might lead to underestimation of the safety effectiveness of installing signals at relatively unsafe intersections. Most of the studies reviewed were conducted before these pitfalls came to light. Nevertheless, the upshot of these revelations is that there is very little substantial knowledge about the safety impact of traffic signal installation. The question of how to improve this somewhat embarrassing state of affairs using the latest methods of analysis is addressed, and recommendations are made on how to incorporate useful knowledge into the allocation of resources for signal installation.

The net impact of traffic signal installation is usually assessed by evaluating the effects on vehicle operating costs, motorist delay, vehicle emissions, and safety. In evaluating the likely safety impact, a traffic engineer would ideally like to draw on the knowledge accumulated in numerous past evaluations. However, as many previous reviewers have noted [see paper by Box and Alroth (*1*), for example], these studies have failed to produce a consensus on the safety impact of signals. Yet these studies are apparently the principal source for many prevailing beliefs about the conditions under which signals are likely to reduce or increase accidents. Even where there is a semblance of consensus, as on the influence of preinstallation accident history, one cannot, as will be contended in this paper, be certain that what is taken as common knowledge is based on a strong foundation.

One important but often overlooked explanation for the shaky foundations of knowledge on this subject is that there might have been practical limitations or methodological deficiencies in several of the studies on which current knowledge is based. Before the results of a study are accepted as useful knowledge, it seems prudent to examine the study to ensure that the assumptions and methods of analysis are proper, that the conclusions drawn are supported by the data, and that the findings are applicable. This is the guiding principle of the review on which this paper is based.

2 Elmbrook Crescent, Etobicoke, Ontario M9C 5B4, Canada.

The paper begins with a brief description of the apparent status of knowledge on a number of issues. The main focus of this paper is to point out how errors of analysis and interpretation committed in many evaluations might cast a shadow on much of the empirical evidence on the safety impact of signal installation. Therefore, a special section is devoted to illustrating two of the pitfalls that plague safety studies. Against this background, the conventional wisdom on the safety impact of signal installation is then scrutinized. Finally, there are some suggestions on how the status of knowledge might be improved.

To provide some background for the discussion to follow, relevant features and results for a selected number of before-and-after studies are summarized in Table 1. (Note that in Table 1 and throughout this paper, percent reductions are specified, so negative values indicate increases in accidents.)

APPARENT STATUS OF KNOWLEDGE

Overall Safety Impact

The data in the fourth through eighth columns of Table 1 indicate an apparent lack of consensus on the overall safety impact of signal installation. This inconsistency is not surprising to most reviewers, who generally recognize that the safety effect of signal installation is likely to depend on a complex web of factors—approach volumes, intersection geometry, signal design, accident history and characteristics, approach speeds, and so on. Yet it is not uncommon to find accident reduction factors for signal installation specified in such authoritative sources as the Institute of Transportation Engineers (ITE) handbook (*15*) (18 percent for total accidents, 32 percent for injuries, and 49 percent for fatalities). On the basis of Table 1, the source or applicability of such factors is unclear.

It should be pointed out that the studies cited in Table 1 are of the before-and-after type. Another type of study is the cross-section type, in which the accident experience of signalized intersections is compared with that of unsignalized intersections. The results of these studies almost always show that signalized intersections have a higher accident rate. As will be discussed later in this paper, it is perhaps incorrect to infer that this higher rate is due to signal installation, because the higher accident rate may have already existed before installation and may actually have been part of the reason for installing the signals.

Effect on Specific Types of Collision

Most traffic engineers are cautious in using factors such as those in the ITE handbook, believing that because signal installation is likely to increase rear-end accidents and reduce right-angle accidents, the overall safety impact should depend on the relative number of each of these two types of accidents. This belief makes good sense intuitively; however, as the data in Table 1 indicate, the support is weak—the impact of signal installation on each of these two accident types is reasonably, but not totally, consistent.

Effect on Accident Severity

Table 1 also challenges some other beliefs. For example, there is a reasonable belief that right-angle accidents are more severe than rear-end accidents; thus by reducing

TABLE 1 PERCENT REDUCTION IN ACCIDENTS AFTER SIGNAL INSTALLATION

Author/ Reference/ Study Period *	Location	No. of sig-nals	% reduction by accident type					
			Total	Rear-End	Right Angle	Injury	Left Turn	Other Issues **
Solomon (2)	Michigan Rural	39	-23	-200	+51	+20	n/a	B,D
King (3) 1-2 yr.	Virginia ?	30	-24	-181	+34	+18	-16	C,D
King (3)	Michigan ?	33	-8	-84	+45	n/a	-236	
N.Y.DOT (4) 3-4 yr.	New York Rural ?	39	+7	+21	+13	-11	-13	D
Hammer (5)	California Rural ?	170	+21	-90	+76	+32	+14	
Clyde (6)	Michigan Urban	52	-34	-98	+45	-11	-66	
Short (7) 3 yr.	Milwaukee Urban	31	+2	-37	+34	-6	n/a	A,B,C,D
Vey (8)	22 cities Mixed	599	+20	-37	+56	n/a	n/a	C
Cribbins (9) 1 yr.	N.Carolina Rural	19	-7	-147	+73	-21	-21	D
Malo (10)	Detroit Urban	20	+47	+24	+75	n/a	n/a	
San Francisco (11) 1 yr.	Urban	48	+53	+72	+80	+50		A,B,C,D
Leckie (12) 1 yr.	Ontario Rural ?	13	+8			-27		
Schoene (13) 2 yr.	Illinois Rural	30	-16	-221	+48	-26		A,B,C
Smith (14) 1 yr.	California Mixed	32	+39					

* Typical lengths for before and after periods, where given.

** A - Safety of warranted vs unwarranted installations
B - Influence of volumes
C - Influence of pre-installation accident record
D - Influence of geometrics, layout, and/or approach speed.

right-angle accidents, signal installation will reduce injury accidents. From Table 1, it is not at all clear that this is so; there is no consensus.

Several researchers have attempted to assign weightings to accidents of different severity, but findings with respect to these weighted values vary from study to study or are inclusive. Abramson (*16*) reported an increase in total accidents after signal installation with no significant change in accidents weighted by severity. Because few details of the study were provided, the methodology could not be scrutinized; thus caution should be exercised in accepting these results. Chung-Cha (*17*) used Abramson's weightings and reported that signalized intersections do not necessarily result in smaller figures of merit than STOP-sign-controlled intersections. As will be discussed in depth later, this type of result could be misleading, because it is based on a cross-section type of analysis. A before-and-after study by Cribbins and Walton (*9*) of 19 low-volume, high-speed (45 mph) rural intersections is particularly relevant in the context of this paper. Although there was a 7 percent increase in total accidents (which was insignificant), the authors found a decrease (insignificant) in accidents "normalized" to equivalent property damage only (EPDO) accidents. At undivided highway intersections, there was a significant 41 percent decrease in EPDO accidents. Though these results are interesting, one cannot readily accept them because there are hints that high-accident locations tended to be selected; as discussed later in this paper, such simple before-and-after comparisons may overestimate safety benefits.

Effect on Pedestrian Accidents

Requests by the public for signal installation are quite common, and these are often inspired by a concern for pedestrian safety. It seems important, therefore, to know how signal installation affects pedestrian accidents. However, in the studies reviewed, pedestrian accidents either were not reported or were so few that the results were inconclusive. In a study by Short et al. (*7*), pedestrian accidents decreased (insignificantly) from 19 in 3 years to 12 in a similar period after 31 signal installations. In a study by the New York State Department of Transportation (NYSDOT) (*4*), these accidents increased from 0.1 per year to 0.2 per year after signal installation. Findings in a study by Yagar (*18*) suggest that installation of signals at 25 metropolitan Toronto intersections that had pedestrian crossovers decreased the pedestrian accident rate by about 40 percent; on the other hand, at the 43 installations that had no prior pedestrian protection, the pedestrian accident rate increased by about 20 percent. These statistics, however, were based on an average of fewer than 30 pedestrian accidents per year, and Yagar points out that the 20 percent increase is insignificant. In addition, it is not clear whether there were dramatic changes in pedestrian counts after signal installation; in particular, a reduction in pedestrian volumes might suggest a "migration" of pedestrian traffic, and possibly pedestrian accidents, to other locations.

Warranted Versus Unwarranted Signals

Most jurisdictions have installation warrants for traffic signals, many of which are based on the U.S. Manual on Uniform Traffic Control Devices (MUTCD) (*19*). Traffic engineers' faith in these warrants is founded in part on a belief that installation of warranted signals is likely to reduce accidents, whereas unwarranted signals can increase accidents (*20*). This belief is apparently a reflection of evidence on the effect of traffic volume or preinstallation accident record, or both; this evidence is summarized next.

Influence of Traffic Volume

Warrants are generally based on minimum traffic volume or accident levels, or both. Because most installations tend to be warranted on the basis of traffic volume, the belief that signals improve safety only where they are warranted is apparently a reflection of another belief—that the safety benefit of signal installation increases with traffic volume. There seems to be a concurrence in many studies that traffic volumes play a role, but there is conflicting evidence on the nature of the influence. To illustrate, in Figure 1, which is based on several well-known studies, the influence of traffic volume is not clear. Because the unwarranted signals would tend to be in the lower volume range, it is unclear how this graph, taken as a whole, would support a belief that installation of unwarranted signals can increase accidents.

The inconsistency in Figure 1 might possibly be explained by evidence indicating that the safety effect of signal installation depends not only on the total entering volume but also on how this volume is divided among the approaches. In this respect, the before-and-after study by Schoene and Michael (*13*) of 30 rural intersections is the classic. They reported that accidents generally increased after signalization of intersections with major:minor volumes greater than 6:1; for warranted installations the cutoff ratio was even lower (4:1). Though these results might be somewhat useful, the small sample size precludes the establishment of a relationship among safety effect, total volume, and ratio of major-road volume to minor-road volume.

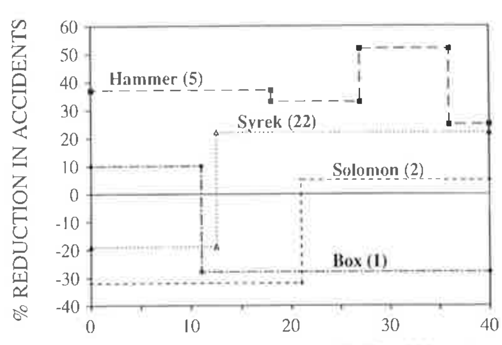

FIGURE 1 Safety impact of signal installation by total entering volume.

That the safety effect of signal installation depends on the details of the traffic volume is underscored in some recent work by Hauer et al., which is published elsewhere in this Record. They explored various models to estimate the safety of a signalized intersection on the basis of information about its traffic flow and accident history and concluded that "logically sound models require that frequency of collisions be related to the traffic flows to which the colliding vehicles belong and not to the sum of entering flows." Because this relation has never been attempted in the examination of the safety effect of signal installation, it would seem that there is a lot yet to be learned about the dependence of this effect on volume. How this might be accomplished is addressed later.

Influence of Preinstallation Accident History

Most jurisdictions permit signals to be warranted, at least in part, on the basis of accident history. For example, many jurisdictions use the MUTCD warrant (*19*), which specifies that if 80 percent of the volume warrant, and some other conditions, is met, a signal may be warranted if five or more "correctable" accidents have occurred in a 12-month period. The Ontario warrant (*23*) is similar, but emphasis is placed on averaging the accident totals over a 36-month period; this would indicate that 15 or more correctable accidents are required over 36 months. In a Canadian warrant (*24*), priority points are assigned that generally increase with, among other factors, the total number of preinstallation reportable accidents per year (averaged over 4 to 5 years). Signal installation is warranted at intersections that have at least 100 priority points. Accident priority points are determined from a chart similar to Figure 2, which actually suggests negative priority points for intersections with fewer than eight accidents per year before signalization. Taken together with the MUTCD and the Ontario accident warrants, the sloping line in Figure 2 must therefore be a reflection of the belief that not only does the safety impact of signal installation increase with the number of preinstallation accidents, but

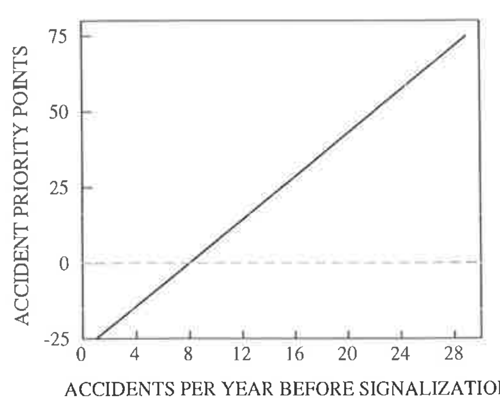

FIGURE 2 Accident priority points applied in Canadian signal warrant (*24*) for $200 minimum for reportable accidents.

also that signals can increase accidents if this number is small enough.

At a glance there appears to be formidable support for this belief. Vey (*8*), in the earliest reported study on this subject, found that installing signals at intersections with three or fewer accidents per year increased accidents by 39 to 70 percent, whereas at other intersections accidents decreased by 19 to 49 percent. In another widely quoted study, Schoene et al. (*13*) indicated an increase in accidents at intersections that had fewer than 10 right-angle accidents for 2 years before signalization. King and Goldblatt (*3*) found that total accidents increased following signal installation at intersections not meeting the MUTCD accident warrant (those with fewer than five right-angle accidents per year). Hakkert and Mahalel (*25*) compared before-and-after accident records for 34 urban signal installations in Israel and found that at intersections with more than five accidents per year, there was generally a decrease in accidents after installation of a signal; at installations with fewer than two accidents per year, accidents generally increased. Presumably on this basis, they concluded that where warranted according to vehicle volumes, traffic signals "most probably have a beneficial effect on the number of accidents."

As will be seen later, much of the evidence on the influence of the preinstallation accident record on the safety impact of traffic signals is tainted by the fact that increases and decreases in number of accidents at intersections may have little to do with changes in safety and a lot to do with random fluctuation; increasing the length of the "before" period to 3 to 5 years reduces, but does not eliminate, this bias. This problem may render useless much of the evidence on this issue, as well as the warrants that seem to be based on this evidence.

Other Evidence on the Impact of Warrants

Some studies do not separate those signals warranted on the basis of traffic volumes and those meeting the safety warrant, and so leave unclear the influence of the individual factors. As will be seen, the validity of these studies is also subject to question. Hanna et al. (*26*) conducted a cross-section analysis of signals already installed in rural areas, found similar accident rates for warranted and unwarranted installations, and inferred that this contradicts common belief that unwarranted signals cause higher accident rates. This inference might be "unwarranted," as will be seen later when the dangers of cross-section studies are discussed. Another study that appeared to contradict the belief about the increase in unsafe conditions after unwarranted installations was the 3-year before-and-after comparison of 31 urban installations by Short et al. (*7*), who concluded that "whether or not a signal was warranted seemed to have little bearing on whether accidents would be reduced," and then claimed, without reporting the evidence, that "one of the best measures found to indicate accident reduction is the average number of correctable accidents per year over a several year period."

In a well-known study, Young (*27*) conducted a one-year before-and-after comparison of 32 urban installations and found that of seven intersections with reductions of more than five accidents, five met at least one MUTCD (*28*) warrant, and three met the accident warrant; conversely, of 10 intersections showing an increase in accidents after signalization, "only 2 met any semblance of a warrant." On this basis, Young concludes that his evaluation shows "a pattern of support for the present warrants with respect to the effect of signalization on safety." However, apparently in contradiction of this statement, Young goes on to state that the study gives very little support to the idea that unwarranted signals per se will increase accidents, because accidents actually decreased at nine unwarranted installations. In any case, as will be described in the next section, this evidence may be almost totally useless.

PITFALLS IN SAFETY EVALUATION

Regression to the Mean

Perhaps the most common pitfall in studies of the safety impact of signal installation is the failure to account for regression to the mean. This cumbersome term refers to a phenomenon whereby, even if signal installation has no safety impact, intersections that recorded many accidents in the period before signalization would, on average, record fewer accidents in the after period, whereas accidents would increase at intersections that had few accidents. By simply comparing before-and-after accident totals or rates, therefore, one could erroneously conclude that signal installation improved safety only where the safety warrant was satisfied. Furthermore, if, as is sometimes the case, intersections are selected for signal installation on the basis of an unusually high accident record, a simple before-and-after comparison is likely to overestimate the effect of the treatment. [For more details, see the paper by Hauer and Persaud (*29*).]

To illustrate, right-angle accidents recorded in two consecutive 2-year periods at unaltered, two-way stop-controlled intersections in Philadelphia are summarized in Table 2. [The data are taken from a thesis by Ebbecke (*30*).] The first row of numbers shows that 17 intersections had no right-angle accidents in either period, 11 that had none in the first period had 1 in the second period, and so on. The numbers connected by the diagonal line are for intersections whose right-angle accident total remained unchanged; the numbers to the right of this line are for intersections at which an increase in right-angle accidents was reported, whereas the number of intersections that recorded a decrease are to the left of the diagonal line.

Summation of the appropriate numbers in Table 2 shows that, even though the intersections remained essentially unaltered, right-angle accidents increased at many intersections (80), decreased at still more (88), and remained unchanged at only 43 intersections. This type of conclusion accords precisely with the common belief that signals sometimes increase and sometimes decrease accidents, yet it was derived from data on intersections that remained unchanged and results from changes largely due to random fluctuation. The upshot of this paradox is that statements about the frequency of increases and decreases in accidents following signalization must be disregarded unless one can determine how much of the change is due to random fluctuation and how much is due to signal installation.

Although the most common method of measuring safety impact is from the changes in the number of accidents or in the accident rate (per unit volume), many studies report their findings and make conclusions on the basis of the number of intersections showing a decrease in accidents and the number showing an increase following signal installation. Among the most notable is the widely quoted study by Young (*27*), who reported that, on signalization, the 5-year accident total increased at 102 intersections, decreased at 23, and did not change "significantly" at 27. Then, noting that it was "much more common for accidents to increase," Young concluded that there was no obvious advantage of signalization as an accident prevention measure. In a later study reported in the same paper, a 1-year before-and-after comparison of 32 subsequent installations found accident reductions to be more prevalent, which led to the conclusion that "this study shows far more favorable results from signalization than did the earlier study." Clearly, the relative size of the increases and decreases should matter, so one should be cautious about accepting this statement. Furthermore, as shown earlier, statements about the number of intersections with increases and decreases are not very meaningful because such changes can be due mainly to random fluctuation.

This phenomenon also appears to be present in the several studies (*13, 14, 27*) that attempt to identify common factors (volume, accident history, warrant compliance) among the intersections with increases (or those with decreases) in accidents. Invariably, as mentioned earlier, these studies find that the intersections with increases in accidents following signalization tend to have had low volumes or few accidents, or both, before signalization. Because these are the intersections that tend not meet the common warrants (e.g., those of the MUTCD), it is often concluded that unwarranted installations can increase accidents. Perhaps the chicken came before the egg, and accident warrants, such as those of the MUTCD, are in part based on the finding that signals improve safety only where there were numerous accidents. This might also be the inspiration for the Canadian warrant (*24*) (Figure 2) that assigns negative priority points to intersections with fewer than eight (total) accidents per year.

An examination of the data in Table 2 shows how one can erroneously conclude that only warranted signals reduce accidents; one can arrive at this conclusion even if signal installation has no safety effect. For example, the 204 intersections with less than 10 right-angle accidents in the first period might not satisfy the MUTCD accident warrant (five or more "correctable" accidents in 12 months). Table 3 shows that these "unwarranted" inter-

TABLE 2 RIGHT-ANGLE ACCIDENTS AT TWO-WAY-STOPPED INTERSECTIONS

Accs. in Period 2 / Period 1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Number of intersections with indicated numbers of accidents in each period														
0	17	11	7	4	1	2	1								
1	10	9	4	5	1	3	1								
2	6	7	5	5	6	3	2	1	1	2					
3	3	10	7	3	3	1		1	2			1			
4	3	2	4	3	5	1			2	1					
5		1	2	1	2	3	1	1		2		1			
6	2			2	1	3									
7		1	3	1	2	1		1							
8		1	1	1	1	1							1		
9														2	
10			1												
11								1	1						
12			1						1						
13															
14											1				
..															
20															1

TABLE 3 CHANGES IN RIGHT-ANGLE ACCIDENTS IN GROUPS OF UNALTERED INTERSECTIONS

GROUP	No. of intersections	Total accidents in period 1st. 2 years	2nd. 2 years	CHANGE
10 or more accidents	7	90	54	Decrease of 36
Less than 10 accs.	204	525	558	Increase of 33

sections had an increase of 33 accidents in the second period; on the other hand, the seven intersections that would satisfy the accident warrant (and be more likely to meet the volume warrant) showed a decrease of 36 accidents (40 percent). So, were the 211 intersections (or a representative sample) to be signalized, and signalization had no effect on safety, one could erroneously conclude that unwarranted, low-volume, low-accident installations increase accidents, whereas warranted, high-accident, high volume installations increase safety. Because this erroneous conclusion is derived from changes in accidents due to random fluctuation, studies that make such claims must be taken with a grain of salt unless, once again, changes due to random fluctuation are extracted.

The problem described in the preceding paragraph can arise whenever the selected group contains a mixture of high- and low-accident intersections. In cases where intersections are selected for signal installation because they have had many right-angle accidents, as might be the case for the seven intersections in Table 2 with 10 or more such accidents, which would meet the MUTCD accident warrant, a further complication arises. As indicated in Table 3, these intersections had a 40 percent decrease in right-angle accidents solely because of random fluctuation. So, were signals to be installed at these intersections, a simple comparison of the number of accidents before and after installation could cause the impact of signalization on right-angle accidents to be overestimated by 40 percent.

It should be pointed out that in cases where there is no selection bias—where selection of intersections for signal installation is not based on accidents recorded at all—the danger of overestimating the overall safety impact would not be present. One can show empirically that a randomly selected group of intersections from Table 2 would show little or no change in the total number of accidents over the two periods. The extent to which intersections are selected for signal installation on the basis of a high number of accidents is not clear, although in the literature there are hints that this is the basis for some installations. For example, the Canadian warrants (*24*), by assigning accident priority points that increase with the total number of accidents, might tend to favor selection of intersections that recorded many accidents. (It should be stated in fairness that this warrant does suggest a 4- or 5-year "before" period; this reduces the overestimation, but it has been shown empirically (*30*) that the overestimation can still be substantial.)

It is also worth emphasizing that even when selection for signal installation is not based on accident record, the danger of regression to the mean still lingers if one were to try to examine the safety effect of subgroups that are identified on the basis of accident record. To illustrate, consider the study by Kay et al. (*31*), who compared accident records before and after signals were removed and found that, on signal removal, accidents increased at intersections that had few accidents and decreased at other intersections. On this basis, the authors suggested that intersections with at least two or more accidents per year are "good candidates for signal removal." For reasons given earlier, this conclusion cannot be made on the basis of a simple before-and-after comparison of intersections grouped according to accident record before the change. The irony in the conclusion by Kay et al. is that the potential danger of regression to the mean was recognized, but the assumption was made that, because they "were not selecting 'extreme' cases for treatment," there was a small chance that the findings would be affected. As Table 4 shows, this phenomenon might, in fact, have substantially affected their findings; right-angle accidents at unaltered signalized intersections in Philadelphia decreased at those with fewer than two accidents per year in one period, whereas there was a compensating decrease in these accidents at other intersections.

The main message from this rather detailed description of one of the common pitfalls in safety evaluation studies is that, in the comparison of accident frequency before and after signal installation, one must separate the changes that are due to random fluctuation from the changes that might be due to signal installation. It should be noted that most of the studies on which our knowledge about the

TABLE 4 CHANGES IN ACCIDENTS AT UNALTERED SIGNALIZED INTERSECTIONS

Accidents in Group	Number of Intersections	Two-year Accident Totals Period 1	Period 2	Change
4 or more	298	2221	2040	8% Decrease of 181 accs.
< 4	79	196	376	92% Increase of 180 accs.

safety impact of signal installation is based were undertaken before the traffic engineering profession came to recognize the dangers of regression to the mean. This partly redeems the studies, but cannot justify the continued acceptance of knowledge based on questionable findings.

Incorrect Inference from Cross-Section Studies

A second major pitfall in safety evaluation occurs when one tries to extend the results of cross-section ("with and without") studies to make inferences about the safety impact of adding a feature. As an illustration, consider the study by Syrek (*22*), who compared accident rates for intersections with different forms of control and found, for example, that signalized intersections with major- and minor-street annual average daily traffic (AADT) of 10,000 and 8,000, respectively, had higher accident rates than STOP-controlled intersections with similar volumes. This might be taken to indicate that installing signals at intersections with such volumes can cause accidents to increase, but what is overlooked in this type of inference is that the signalized intersections might have had a higher level of unsafe conditions for a variety of reasons that may have nothing to do with the installation of signals. Indeed, this higher level of unsafe conditions might have been partly the cause for the signal installation rather than its effect.

In a similar type of study, David and Norman (*32*) examined accidents during 3 years at intersections with various forms of control and found that STOP-controlled intersections had 30 to 60 percent fewer accidents than signalized intersections with similar volumes and the same number of approaches. They also found that two-phase intersections with left-turn lanes had a higher accident rate than those without, and inferred that the introduction of left-turn lanes increases accidents. The inference might be incorrect because the difference in safety might have existed before the introduction of left-turn lanes and might have been partly the cause for their introduction. Similarly, one might question the basis for the recommendation by David and Norman that the introduction of a third phase will reduce accidents at those signalized intersections already having left-turn lanes. Although the subject of this paper is not the safety impact of turn lanes or signal phasing, this digression was made to point out that, because statements about the safety impact of turn lancs and left-turn phasing were made on the basis of a cross-section study, it would not be inconsistent to make the absurd inference from the results of the same study that signal removal decreases accidents by 30 to 60 percent, which therefore makes it a worthwhile countermeasure.

The results of cross-section studies often differ from the results of before-and-after studies. Hanna et al. (*26*) found that warranted and unwarranted rural installations had similar accident rates and inferred that the safety impact of installation was similar for the two groups; on this basis they suggested that "the need to implement a policy of eliminating unwarranted signals is perhaps not so urgent as in urban areas." They also found that, for a given average daily traffic, signalized intersections had a 29 percent higher accident rate than intersections with STOP or YIELD control, and implied that this meant that signalization can increase accidents. Clearly, the inferences by Hanna et al. should be interpreted with caution, as were those of David et al., because differences in safety may have existed before the installation of signals and may have had little to do with the installation of the signals. Indeed, the 21 of 76 installations in the study by Hanna et al. that did not meet the volume warrants might have been justified by a concern for high accident rates—rates higher than those at intersections that remained STOP or YIELD controlled.

Perhaps the most telling example of the differences between cross-section and before-and-after studies is provided by studies by Thorpe (*33*) and Andreassend (*34*), which used essentially the same data base (accidents in Melbourne, Australia). Thorpe, in a cross-section study, found that signalized and unsignalized intersections had similar accident rates, whereas Andreassend, in a before-and-after study, found that signalization was followed by a "highly significant" 32 percent reduction in accidents. If in fact the signalized intersections had higher preinstallation accident rates than the other intersections, Andreassend's before-and-after comparison might have overestimated the safety benefit (because of regression to the mean); on the other hand, inferring from Thorpe's results that signals have no safety effect would underestimate the safety benefit, because such an inference would incorrectly assume that the preinstallation accident rate for signalized intersections was similar to that for other intersections.

To summarize the discussion of cross-section type studies, it seems clear that inferences from this type of analysis about the safety impact of installing signals must be interpreted with caution. Attributing differences in rates between intersections with and without signals to the installation of signals will almost certainly underestimate the safety effectiveness.

HOW TO IMPROVE THE STATUS OF KNOWLEDGE

In the previous sections of this paper, it has been suggested that very little is known about the safety impact of traffic signals and that the accident warrants for traffic signals are, despite appearances, not well supported. In suggesting how this state of affairs can be improved, it is perhaps best to first consider what tools should be available to jurisdictions responsible for decisions on signal installation. These might be described as follows:

1. A method for quantifying the likely safety impact of a contemplated installation,
2. Classification of circumstances under which signal installation is likely to be good or bad for safety, and

3. The incorporation of knowledge gained with the second tool into signal warrants or perhaps into a cost-benefit–resource allocation procedure for signals.

Suggestions on how these tools might be developed follow.

Safety Impact of Signal Installation (Tools 1 and 2)

To provide these tools, it is first necessary that some resource be undertaken to provide estimates of the safety impact of signals already installed and to relate these estimates to various installation circumstances—traffic volumes, the preinstallation accident experience and control type, geometric characteristics, approach speeds, signal design elements, and so on. The outlines of a suggested approach are described below.

To avoid the pitfalls of earlier studies, it is necessary to recognize the distinction between the accident count at an intersection for a short period of 1 to 5 years and its safety—defined as the number of accidents per year expected to occur in the long run at the intersection. We have already seen that comparing accident counts before and after signal installation can produce misleading results about the conditions under which signal installation is likely to be good or bad for safety. This bias would be present even if before-and-after periods of as long as 5 years are used, but it can be eliminated if one compares the long-run expected number of accidents per year (safety) before and after installation. The problem at hand therefore amounts to (a) estimating this quantity defined as safety for signalized and unsignalized intersections and (b) for actual installations, estimating the changes in safety and relating these estimates to various installation circumstances. How to provide these two types of estimates is addressed separately.

Estimating Intersection Safety

The theory for estimating safety is quite recent but is already in place and has been applied to rail–highway grade crossings (*35*) and to signalized intersections (Hauer et al., paper elsewhere in this Record). One would now like to be able to provide an estimate of the safety of an intersection before it has been signalized. This will be indicated by the expected (long-term average) number of accidents of various types for that intersection. These intersections might be categorized by number of approaches and approach lanes, by type of control (two-way or four-way stop controlled), and by location (urban or rural).

To apply the method, detailed data are required for at least 100 intersections in each category of traffic volumes (individual turning movements) and accidents (to the level of movements of vehicles before impact). [Recent work by Hauer et al. (paper elsewhere in this Record) has shown that the best predictor of collisions is the individual vehicle movements related to given types of collisions, and not measures such as total entering volumes.] For each intersection category and each combination of precollision movements, regression analyses could be performed to estimate the expected number of accidents as a function of the relevant volumes. This analysis provides equations for each category of intersection and collision type of the following form:

$$E(m) = \text{function (volume)}$$

where $E(m)$ is the expected number of accidents per year of a given type for a particular category of intersection. Now, in each category, some intersections will actually be safer than others with the same value. The reason for this is that, in grouping intersections, one could never obtain perfect homogeneity. It is therefore necessary to refine $E(m)$ for individual intersections in a category. This can be done by using the accident count (x) at an intersection in some before period. It has been shown (*35*) that, under reasonable assumptions,

$$\begin{aligned} E(m \mid x) &= \text{function}[E(m),x] \\ &= x + [E(m)/\text{Var}(x)]\,[E(m) - x] \end{aligned}$$

This is the estimate of safety that we would like to use in comparing the safety of intersections before and after signal installation. These equations have been developed for signalized intersections in an urban area (Hauer et al., paper elsewhere in this Record). It remains to develop them for other types of signalized intersections and for unsignalized intersections. One can then use them to estimate the safety impact of signal installation.

Estimating the Safety Impact of Signal Installation

At least 2 years of detailed traffic and accident data before and after as many installations as possible (a minimum of 100) is required. For each intersection, one could use the procedure developed in the previous section to estimate the expected number of accidents (by collision type) for the intersection before signalization. One could then use equations already developed by Hauer et al. to estimate the expected number of accidents (of given types) after signalization. (These equations are for urban signalized intersections; equations need to be developed for rural installations.) The percent change is calculated and can then be used as the dependent variable in a regression analysis that will use, as independent variables, relevant traffic volumes, the expected number of accidents before signalization, the type of signal, number of approaches and approach lanes, and so on.

Although the development of the procedures seems quite complicated, much of the groundwork has already been laid. By contrast, once the procedures have been translated into the form of interactive personal computer software, they will be quite easy to use.

Incorporating Knowledge on Signal Safety Impact into Warrants (Tool 3)

It has been suggested on the basis of the review detailed in this paper that safety warrants for signalized intersections are on a shaky foundation and that the belief that only warranted signals improve safety might be incorrect. The procedures suggested will not only offer a convenient method of estimating the likely safety impact of a proposed installation, but would also provide the means for rationally weighing the safety impact against other impacts, in effect, for combining the various warrants. In principle this would be similar to the Canadian procedure (*24*), except that with better knowledge about the safety impact, one can now confidently assign dollar values to the various effects instead of priority points. These procedures lend themselves well to implementation on a personal computer, which would eliminate the need for simplified charts such as Figure 2.

It should be emphasized that, even with current procedures for deciding whether a signal is warranted, it is still necessary to know about the likely safety impact of an installation and to decide how this knowledge will be used. One could, for example, have a policy that a signal is not warranted if it will increase accidents or that a signal is warranted if it satisfies 80 percent of the volume warrant and saves at least one accident per year.

SUMMARY AND CONCLUSIONS

The original intent of this paper was to present a balanced review of knowledge about the safety impact of signal installation. The potential defects in this knowledge turned out to be so universal and so serious that it was inevitable that what started out as an attempt to separate what is useful from what is not resulted in a paper with a distinctly negative tone.

Most will agree that a single set of accident reduction factors for signal installation is meaningless, because the safety impact is likely to depend on a complex web of factors. However, little or nothing is known, or there is no consensus, about the influence of important factors such as intersection geometry, volume, and the type of signal design.

Two potentially damaging pitfalls in safety evaluation studies were described. It was shown that these pitfalls might cast a shadow of doubt on much of the knowledge about the safety impact of signal installation. In particular, the foundation of the belief that, where unwarranted, signal installation is likely to increase accidents appears to be very shaky.

Much of our knowledge on this subject comes from studies that were conducted before these pitfalls and methods of avoiding them (Hauer et al., paper elsewhere in this Record; *29, 35, 36*) were known to researchers. This redeems the early researchers but cannot justify the continued adherence to beliefs that might be based on a shaky foundation. To restore respectability to the status of knowledge on this subject, it is necessary to take the drastic step of starting afresh and use recent data (assuming that the quality of data has improved over the years) and the much more advanced methods of analysis now available to come up with reasonably accurate answers to the following question: Given a set of circumstances for an intersection (approach volumes and speeds, geometry, accident history, signal design, and so on), what is the expected safety impact of installing a signal?

ACKNOWLEDGMENT

This paper is based on research performed under a contract with the Ontario Ministry of Transportation, whose support is gratefully acknowledged. The contents are the sole responsibility of the author.

REFERENCES

1. P. C. Box and W. A. Alroth. Assembly, Analysis and Application of Data on Warrants for Traffic Control Signals, Part III. *Traffic Engineering,* Vol. 56, 1968, pp. 14–21.
2. D. Solomon. Traffic Signals and Accidents in Michigan. *Public Roads,* Vol. 30, No. 10, 1959, pp. 234–237.
3. G. F. King and R. Goldblatt. Relationship of Accident Patterns to Type of Intersection Control. In *Transportation Research Record 540,* TRB, National Research Council, Washington, D.C., 1975, pp. 1–12.
4. *Effect of Signalization.* Traffic and Safety Division, New York State Department of Transportation, 1982.
5. C. G. Hammer. *Evaluation of Minor Highway Improvements,* Part 77: *Traffic Signals.* California Division of Highways, Sacramento, 1970.
6. M. N. Clyde. Michigan Study Indicates Signals Increase Accidents. *Traffic Engineering,* Vol. 35, No. 2, 1964, p. 32.
7. M. S. Short, G. A. Woelfe, and C. J. Chang. Effects of Traffic Signal Installation on Accidents. *Accident Analysis and Prevention,* Vol. 14, No. 2, 1982, pp. 135–145.
8. A. H. Vey. Effect of Signalization on Motor Vehicle Accident Experience. *Proceedings, Institute of Traffic Engineering,* 1933.
9. P. Cribbins and J. Walton. Traffic Signals and Overhead Flashers at Rural Intersections: Their Effectiveness in Reducing Accidents. In *Highway Research Record 325,* HRB, National Research Council, Washington, D.C., 1970, pp. 1–14.
10. A. F. Malo. Signal Modernization. In *Special Report 93: Improved Street Utilization Through Traffic Engineering,* HRB, National Research Council, Washington, D.C., 1967, pp. 96–116.
11. *Study of High Accident Intersections.* San Francisco Department of Public Works, 1974.
12. D. A. Leckie. *Traffic Signalization—Before and After Accident Analysis.* Department of Highways, Ontario, Canada, 1971.
13. G.W. Schoene and H. L. Michael. *Effects of a Change in the Control Device on Intersection Accidents.* Engineering Reprint CE 237. Purdue University, West Lafayette, Ind., 1968.
14. W. C. Smith and J. J. Vostrez. Evaluation of Minor Highway Improvement Projects. *Highway Research News,* No. 13, June 1964, pp. 21–29.
15. W. S. Homburger, L. E. Keefer, and W. R. McGrath. *Transportation and Traffic Engineering Handbook.* Institute of Transportation Engineers, Washington, D.C., 1982.

16. P. Abramson. An Accident Evaluation Analysis. In *Transportation Research Record 486,* TRB, National Research Council, Washington, D.C., 1974, pp. 33–41.
17. Chung-Cha Liu. *Some Effects of Unwarranted Traffic Signals.* Joint Highway Research Project IN/JHRP-79–13. Purdue University, West Lafayette, Ind., 1979.
18. S. Yagar. *Safety Effects of Converting Intersections in Metro Toronto to Pedestrian Crossovers or Traffic Signal Control.* Metropolitan Toronto Roads and Traffic Department, Toronto, Ontario, Canada, 1984.
19. *Manual on Uniform Traffic Control Devices for Streets and Highways.* Federal Highway Administration, U.S. Department of Transportation, 1978.
20. Roy Jorgenson and Associates, Inc. *NCHRP Report 197: Cost and Safety Effectiveness of Highway Design Elements.* TRB, National Research Council, Washington, D.C., 1978.
21. P. C. Box. *Traffic Control and Roadway Elements—Their Relationship to Highway Safety,* Chapter 4: *Intersections.* Highway Users Foundation for Safety and Mobility, Washington, D.C., 1970.
22. D. Syrek. Accident Rates at Intersections. *Traffic Engineering,* Vol. 25, No. 8, 1955, pp. 312–316.
23. *Traffic Signals,* Division 2: *Installation Warrants for Traffic Signals.* Ontario Ministry of Transportation, 1982.
24. *Uniform Traffic Control Devices for Canada,* 3rd ed. Part B, Division 2: *Installation Warrants for Traffic Control Signals.* Roads and Transportation Association of Canada (RTAC), Ottawa, Ontario, 1976.
25. A. S. Hakkert and D. Mahalel. Effect of Traffic Signals on Road Accidents—With Special Referenee to the Introduction of a Blinking Green Phase. *Traffic Engineering and Control,* Vol. 19, No. 5, 1978, pp. 212–215.
26. J. T. Hanna, T. E. Flynn, and W. L. Tyler. Characteristics of Intersection Accidents in Rural Municipalities. In *Transportation Research Record 601,* TRB, National Research Council, Washington, D.C., 1976, pp. 79–82.
27. T. E. Young. New Traffic Signals—Their Effect on Street Utilization. In *Special Report 93: Improved Street Utilization Through Traffic Engineering,* HRB, National Research Council, Washington, D.C., 1967, pp. 84–95.
28. *Manual on Uniform Traffic Control Devices for Streets and Highways.* FHWA, U.S. Department of Transportation, 1961.
29. E. Hauer and B. N. Persaud. A Common Bias in Before-and-After Accident Comparisons and Its Elimination. In *Transportation Research Record 905,* TRB, National Research Council, Washington, D.C., 1983, pp. 164–174.
30. G. M. Ebbecke. An Examination of the Areawide Effects of Traffic Control Device Installations in a Dense Urban Area. Master of Civil Engineering thesis. Villanova University, Philadelphia, Pa., 1976.
31. J. L. Kay, L. G. Neudorff, and F. A. Wagner. *Criteria for Removing Traffic Signals.* Technical Report FHWA/RD-80/104. FHWA, U.S. Department of Transportation, 1980.
32. N. A. David and J. R. Norman. *Motor Vehicle Accidents in Relation to Geometric and Traffic Features of Highway Intersections.* FHWA, U.S. Department of Transportation, 1975.
33. J. C. Thorpe. Accident Rates at Signalized Intersections. *Proc., 4th Conference of Australian Road Research Board,* Vol. 4, Part 1, 1968, pp. 995–1004.
34. D. C. Andreassend. Another Look at Traffic Signals and Accidents. *Proc., 5th Conference of Australian Road Research Board,* Vol. 5, Part 3, 1970, pp. 304–318.
35. E. Hauer and B. N. Persaud. How to Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices. In *Transportation Research Record 1114,* TRB, National Research Council, Washington, D.C., 1987, pp. 131–140.
36. E. Hauer, J. Lovell, and B. N. Persaud. *New Directions for Learning about Safety Effectiveness.* Report FHWA/RD-86-015. FHWA, U.S. Department of Transportation, 1986.

Estimation of Safety at Signalized Intersections

EZRA HAUER, JERRY C.N. NG, AND JANE LOVELL

Models to estimate the safety of a signalized intersection on the basis of information about its traffic flow and accident history are provided. They are based on data from 145 intersections in Metropolitan Toronto. Several insights were obtained during the development of the models. First, logically sound models require that the frequency of collisions be related to the traffic flows to which the colliding vehicles belong and not to the sum of the entering flows. Second, it is therefore necessary to categorize collisions by the movement of the vehicles before the collision and not by the initial impact type, as is customary. Third, the relationship between collision frequency and the related traffic flows is at times unexpected in form. For each of the 15 accident patterns, an equation is given to estimate the expected number of accidents and the variance using the relevant traffic flows. When data about past accidents are available, estimates based on traffic flow are revised with a simple equation. Several practical questions can now be answered. Given the traffic flow for a signalized intersection, one can predict how many and what kinds of accidents should be expected to occur on it; one can also show the probability density function (pdf) of the estimate. Knowledge of the pdf allows the determination of what an unusually high number of accidents would be on such an intersection. If the traffic flow of the intersection changes from year to year, one can estimate what changes in safety should be attributed to changes in flow. Also, one can correctly compare the safety of several intersections that have different flow patterns. Most important, one can estimate safety when both flows and accident history are given and, on this basis, judge whether an intersection is unusually hazardous. This method of estimation is recommended for accident warrants in the Manual on Uniform Traffic Control Devices.

In this paper we give equations to estimate the safety of a signalized intersection when the vehicular flows using it are known. This kind of estimate describes the safety of an "average" signalized intersection with given flows. How the safety of such an average signalized intersection depends on the details of the traffic flow may be regarded as basic knowledge, the raw material on which engineering design and decisions are or should be based. Ours is not the first attempt to explore this question, and earlier work will be reviewed first.

The ability to estimate intersection safety as a function of traffic flows may be useful when one has to judge whether the pattern of accidents and their number at some specific intersection are similar to what one might normally expect with such flows. Surely if one wishes to identify what is unduly hazardous and to diagnose what the reasons for such deviation might be, one has to have a good idea about what is normal. To judge what is normal and what is deviant, one has to know also what variability is found in the population of similar intersections. This is why an estimate of the variance of safety will also be provided.

Another circumstance in which one has to know how safety depends on traffic flows is when one has to judge whether some intervention has affected safety and what the extent of the effect is. In this case one has to separate those changes in safety that are due to the intervention from those that are due to the concurrent and inevitable changes in traffic flow.

When for a specific intersection the number of accidents is known (in addition to the traffic flow), this added information must also be reflected in the estimate of safety for this specific intersection. We will show how to combine the safety estimate based on flows with accident data. In our view, it is this combined estimate that should be the basis of the accident warrants used in the Manual on Uniform Traffic Control Devices.

We spoke of "safety" without stating what the word means. "Safety" is the property of some specific entity, in this case that of a signalized intersection. The "safety property" of an intersection is defined as the number of accidents and their adverse consequences expected to occur on it per unit of time. The term "expected" is equivalent to the average in the long run if it were possible to freeze all prevailing conditions that affect safety, such as traffic, weather, driver characteristics, and so on. The safety of some intersection will be denoted m. The mean of the m's in a population of intersections will be denoted $E\{m\}$ and their variance $\mathrm{Var}(m)$.

PREVIOUS WORK

Numerous relationships between accident frequencies and traffic flows have been suggested over the years. A comprehensive survey of these relationships has been given by Chapman (*1*) and Satterthwaite (*2*). The following summarizes findings that pertain to intersections.

Transport Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, M5S IA4, Canada.

Thorpe (*3*), Smith (*4*), and Worsey (*5*) suggest that the number of all accidents at an intersection is proportional to the sum of flows that enter the intersection. The merit of this approach is its simplicity. Its shortcoming is that it is logically unsatisfactory and not a suitable basis for the engineering analysis, which attempts to link cause and effect. One expects, for example, that the number of rear-end accidents at an intersection approach will strongly depend on the flow on approach A and depend only weakly on the flow on approaches B, C, and D. Similarly one should expect that collisions between vehicles from streams A and B moving at right angles to each other might be related to the product of flows A and B. To use the sum of flows A and B leads to the logical difficulty that one will predict accident occurrence even when one of the flows is zero; to use A + B + C + D just compounds two logical difficulties. When a model rests mainly on considerations of correlation and has obvious logical faults, its prediction performance is usually poor. To illustrate, suppose that apples and cherries are grown in the orchards of Ontario. The bushels of fruit produced in Ontario are strongly correlated with the acres of orchards in Ontario and could be estimated on that basis. However, most would agree that it would be better to estimate fruit production by first estimating bushels of cherries on the basis of the acres of cherry orchards and bushels of apples on the basis of the acres planted with apple trees and only then add the two numbers. Not only does one get more detailed information (about apples and cherries, rather than "fruit") but one rightly expects the result to be more reliable. This is the approach that we advocate.

Breunning and Bone (*6*), Surti (*7*), and Hakkert and Mahalel (*8*) relate accidents to the products of the conflicting flows. This is based on the speculation that were drivers blind, the number of collisions could be expected to be proportional to the product of vehicle flows. However, in other empirical research (*9, 10*) it has been found that the number of accidents is in fact not proportional to the product of flows. Rather, accidents were found to be related to the product of flows with each flow raised to a power of less than 1. Tanner (*11*) suggested that the square root of the product of flows would be sufficiently accurate as a rule of thumb.

Support for the "product-of-flows-to-power" relationship can be found in other circumstances as well. For example, when the expected number of accidents at a highway-rail grade crossing is calculated, train flow and highway traffic flow usually enter into the product with an exponent of 0.3 to 0.6 (*12–14*). Similarly, when estimating the expected number of "opposite direction" accidents on two-lane rural roads, Zegeer et al. (*15*) opt for the product-of-flows-to-power model.

Our effort to relate accidents at signalized intersections to traffic flow will be guided by the primacy of logical requirements. First, we will attempt to relate accidents to the traffic flows to which the colliding vehicles belong. This means that accidents between vehicles proceeding in the same direction have to be estimated separately from accidents between, say, vehicles turning left and those proceeding straight through the intersection. (This is analogous to the separation of apples and cherries.) Second, we will examine the data to see what functional relationship is indicated before we decide on the functional form.

It should be obvious that this kind of estimation requires information about turning flows. This puts a strain on practicality when only approach volumes are counted by automatic counters. Even in this case, the suggested estimation procedure can still be used except that turning flows have to be estimated first. Methods for doing so are easily available and commonly used [see, e.g., paper by Hauer et al. (*16*)].

THE DATA

To enhance the chances of success we have selected for analysis a set of signalized intersections that are similar in most respects except traffic flows and accident history. Thus, the data are for 145 four-legged, fixed-time, signalized intersections in Metropolitan Toronto that carry two-way traffic on all approaches and have no turn restrictions. Most are on straight, level sites with a speed limit of 60 km/hr (35 mph).

One-day vehicle counts were collected manually. Thus, for each approach we have details of turning flow and straight-through flow for the a.m. peak, p.m. peak, and off-peak. All vehicle counts are for weekday conditions, and the majority of counting was done during 1984.

The accident data are for 1982, 1983, and 1984. They were derived from the computerized version of the police accident report. A consistency check was performed with the computerized data, but the hard copy of the form was not consulted. To correspond to the available traffic count information, we used only "daytime" accidents, that is, those that occurred between 7:00 and 9:00 a.m. (the morning peak), those that occurred between 4:00 and 6:00 p.m. (the evening peak), and those in the time interval between 10:00 a.m. and 3:00 p.m. (the off-peak). The hours 9:00 to 10:00 a.m. and 3:00 to 4:00 p.m. were excluded from the analysis because this is when signal timing plans change. Accident data were divided into collisions involving pedestrians, single-vehicle accidents, collisions between two vehicles, and collisions involving more than two vehicles. The frequency with which these four classes arise is as follows:

Accident Type	*No.*
Single vehicle	54
Two vehicles	2,084
Multivehicle	248
Vehicle-pedestrian	187
Total	2,573

For the purpose of this analysis, only collisions involving two vehicles were examined. These accidents represent 81 percent of the total accidents.

In summary, the following analysis concerns weekday, daytime crashes of two vehicles at simple signalized intersections.

ANALYSIS OF ACCIDENTS BY PATTERN

To begin the statistical analysis from a logically satisfactory foundation, we thought it best to relate accidents to the flows to which the two colliding vehicles belong. The real interactions are perhaps more complex. However, as a point of departure and in view of the paucity of our data, it appears sensible to begin with the obvious.

The 15 patterns in which two vehicles at a four-legged intersection can collide are shown in Figure 1. Accidents in each pattern are defined by the maneuvers of the two vehicles before the collision (which are recorded on the police accident report as "turning left," "turning right," "going ahead," etc.). Thus, for example, in Pattern 1 the collision is between vehicles that are proceeding straight through the intersection, the crash occurring before the stop line; in Pattern 2 the same vehicles are involved but the crash is within the intersection; in Pattern 13, one of the vehicles is turning right.

We avoid using the more common categorization by initial impact type (such as rear-end, angle, turning movement, sideswipe, etc.) because of its ambiguity. It is indeed illuminating to examine the cross-tabulation of accidents classified by both criteria as is shown in Table 1.

Categorization by initial impact type is most common in the practice of intersection safety analysis. To demonstrate the imprecision of this type of analysis, Table 1 reveals that less than half of the angle collisions arise from Pattern 4; Patterns 5 through 12 all involve a left-turning vehicle, yet only 423 of 910 are classified as turning-movement accidents according to the initial impact type; most approaching initial impacts are not in Pattern 3 but in Pattern 6; and so on. We are led to the conclusion that (as many practicing traffic engineers have known all along) when accidents are categorized by initial impact type, their cause-and-effect relationship with traffic flow is weakened. To understand and analyze accidents at intersections, it is better to use the "vehicle maneuver" entry from the police accident report.

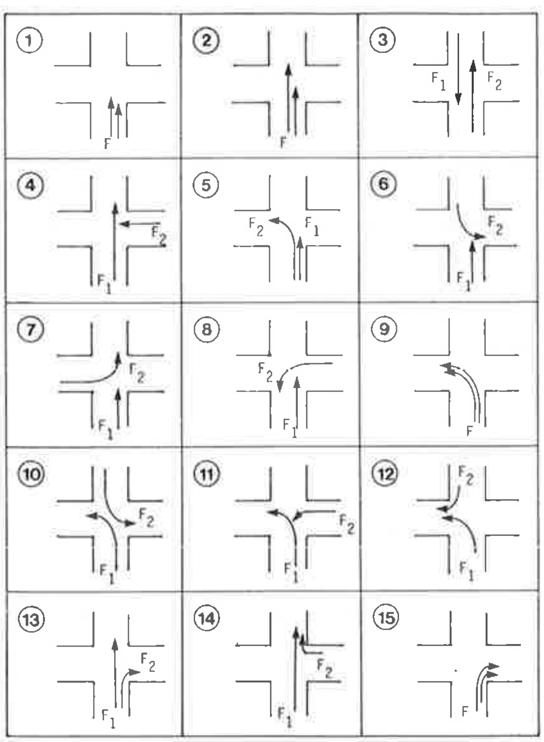

FIGURE 1 Accident patterns by vehicle streams.

EXPLORATION: HOW ACCIDENTS DEPEND ON THE CONTRIBUTORY TRAFFIC FLOWS

To illustrate how the form of the model equations was chosen, consider the typical left-turn accident represented by Pattern 6. For each intersection this pattern arises four times. Therefore, the 145 intersections give 580 "sites," each with different values for the two "contributory" vehicle flows. Table 2 gives the average number of accidents per site in 3 years for five ranges of the left-turning flow versus five ranges of the straight-through flow (in vehicles per day). The cells give the average number of accidents per site in 3 years and, in parentheses, the number of Pattern 6 accidents and the number of sites with flows in the range for that particular cell. The rightmost column and the bottom row give the same information summed over the corresponding row or column. The somewhat irregular flow ranges were selected so that each row and column would have approximately one-fifth of all accidents.

In principle, one should seek a function form that fits the cell entries. However, in this case, the cell entries in Table 2 are based on too few accidents to allow for finesse. Therefore, useful clues are derived from examining the row and column totals. This is shown in Figure 2a and b.

There is a clear suggestion in Figure 2a that within the range of the data, accidents are proportional to the through traffic. Therefore, denoting the through flow in Pattern 6 F_1, we wish to select a model form such that the expected number of Pattern 6 accidents ($E\{m_6\}$) is proportional to F_1. The increase of accidents with the left-turning flow (F_2) appears to be nonlinear. The kind of increase indicated in Figure 2b can be captured by a function such as $F_2^{b_2}$, where b_2 is a coefficient smaller than 1. Therefore, for Pattern 6 accidents, a simple functional relationship that can closely match what we observe in Figure 2a and b is

$$E\{m_6\} = b_0 \times F_1 \times F_2^{b_2} \quad (1)$$

No cause-and-effect arguments were used in selecting this functional form; the guiding principle was the wish to ensure a satisfactory fit with parsimony of parameters and without violation of the obvious logical requirements. We also explored alternative functional forms, which all increased the number of parameters. In no case did this seem worthwhile in terms of the improved fit to the data.

A similar exploration of how Pattern 1 and Pattern 2 accidents depend on traffic flow led to the expected conclusion: in this case, a straight-line fit seems satisfactory.

The examination of Pattern 4 proved more interesting. Because of the symmetry in the situation (see Figure 1) we

TABLE 1 CROSS-TABULATIONS OF ACCIDENTS BY VEHICLE MANEUVER PATTERN AND INITIAL IMPACT TYPE

PATTERN	NUMBER OF ACCIDENTS	INITIAL IMPACT TYPE					
		RE	An	Tn	Ss	Ap	Other
1	538	429	45	5	52	1	6
2	197	109	30	1	52	0	5
3	12	0	4	0	3	5	0
4	258	1	237	0	6	11	3
5	47	10	13	18	5	1	0
6	729	0	329	355	6	32	7
7	42	5	22	14	0	0	1
8	32	0	17	9	1	5	0
9	29	15	0	10	3	0	1
10	2	0	0	2	0	0	0
11	7	1	3	3	0	0	0
12	22	0	7	12	2	0	1
13	73	20	14	29	5	0	5
14	80	5	33	34	8	0	0
15	16	8	3	2	3	0	0
Total	2084	603	727	494	146	55	29

Initial Impact Type :
RE : Rear-End
An : Angle
Tn : Turning Movement
Ss : Sideswipe
Ap : Approaching

TABLE 2 CROSS-TABULATIONS OF ACCIDENTS BY VEHICLE FLOWS FOR PATTERN 6

STRAIGHT-THROUGH FLOW (veh./day)	LEFT-TURNING FLOW (veh./day)					ROW TOTAL
	0-521	522-821	822-1033	1034-1408	1409-4038	
0-3525	0.49 (51/104)	0.60 (35/ 58)	0.63 (20/ 21)	1.57 (33/ 21)	0.56 (9/ 16)	0.64 (148/231)
3526-4825	0.61 (27/ 44)	1.00 (25/ 25)	1.91 (44/ 23)	1.24 (21/ 17)	1.75 (28/ 16)	1.16 (145/125)
4826-5941	1.27 (19/ 15)	1.22 (33/ 27)	1.44 (23/ 16)	1.36 (15/ 11)	2.79 (53/ 19)	1.63 (143/ 88)
5942-7771	1.31 (17/ 13)	1.40 (35/ 25)	3.10 (21/ 8)	2.20 (36/ 8)	1.40 (36/ 11)	1.78 (145/ 53)
7772-12091	1.67 (30/ 18)	2.75 (22/ 8)	2.63 (21/ 8)	4.50 (36/ 8)	3.27 (36/ 11)	2.73 (145/ 53)
COLUMN TOTAL	0.74 (144/194)	1.05 (150/143)	1.56 (139/ 89)	1.93 (149/ 77)	1.91 (147/ 77)	1.26 (729/580)

have chosen to distinguish the two flows by calling the larger flow F_1 and the smaller F_2. In Figure 3a and b we show the observed relationships between the average number of accidents per site for 3 years and the two flows. It appears that, for the range of flows for which data are available, the larger of the two flows exerts little influence on the number of accidents. The smaller of the two flows exerts a great deal of influence initially, but this tapers off as the flows become larger.

This is an unexpected and tantalizing finding. With hindsight one can offer speculative explanations. It is possible, for example, that accidents of this kind involve mostly platoon leaders; the number of vehicles arriving after the platoon leader would in this case be immaterial. Alternatively, one could think of the points in Figure 3a as being a continuation of the points in Figure 3b—the same functional relationship for both flows. Of course, speculation is not a substitute for explanation and does not amount to "understanding."

Thus, on the basis of an exploratory analysis, one can suggest functional forms for expressions that fit what has been observed in Patterns 1, 2, 4, and 6 with parsimony. There are not enough data for any of the remaining accident patterns to warrant similar exploratory analyses. For these we had to select model forms by analogy and judgment.

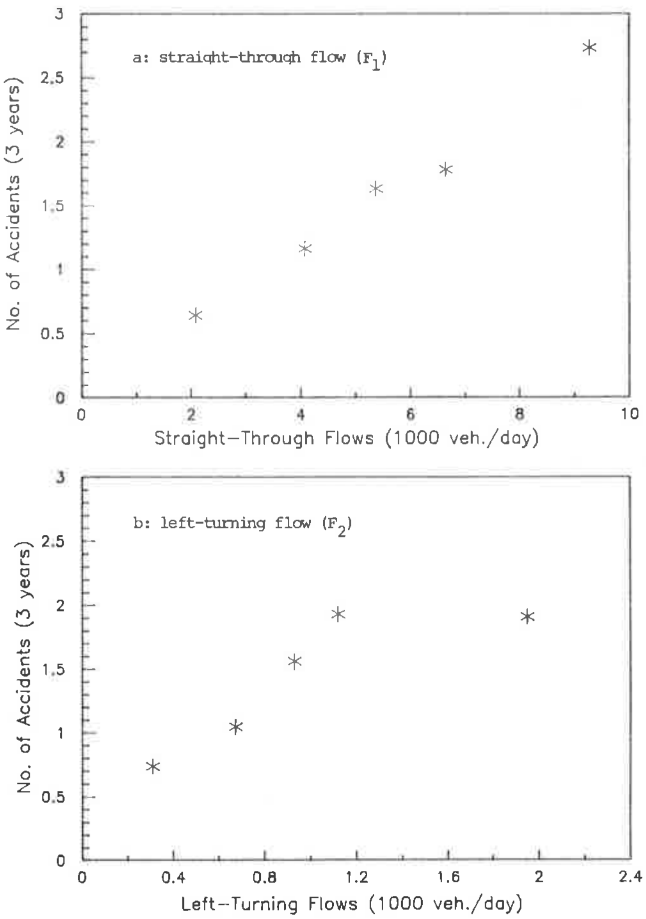

FIGURE 2 Relationship between accidents and contributory flows for Pattern 6.

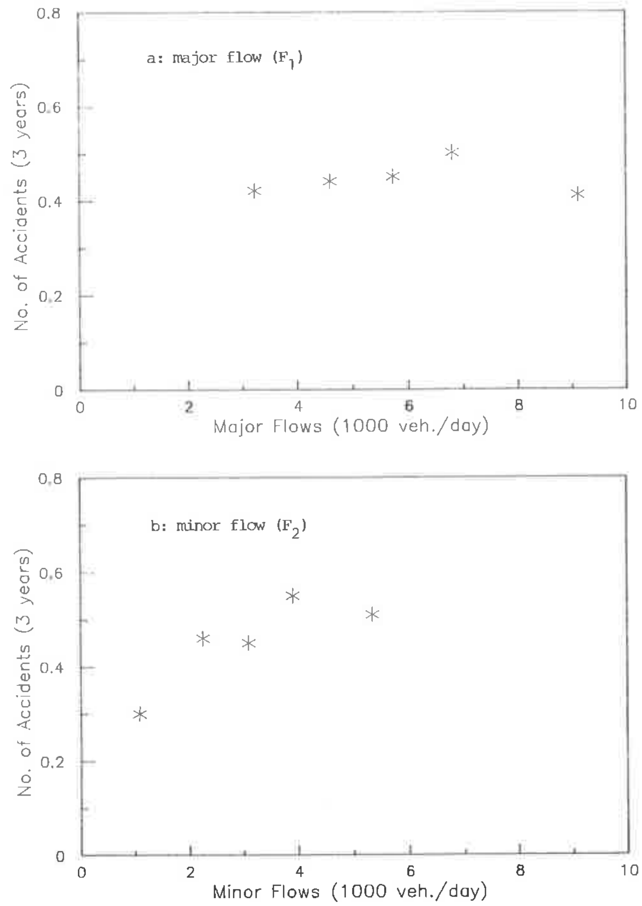

FIGURE 3 Relationship between accidents and contributory flows for Pattern 4.

On the basis of insight gained in the exploration stage we proceeded to the task of examining the statistical performance of the several model forms that appeared promising.

ESTIMATION OF COEFFICIENTS

Coefficient estimation is the domain of the professional statistician. It is a domain too frequently invaded by those who mistakenly believe that a statistical software package can be used by nonexperts. In this paper we do not attempt a comprehensive coverage of technical detail. Rather, we will point out some matters of method that seem important when it comes to the statistical treatment of accident data. It is best to do so with reference to the models at hand.

For accident Patterns 1 and 2 we have concluded that a simple model of the form $E\{m\} = b_0 * F$ seems sufficient. The reflex inclination is to "run a regression" in order to find an estimate $\hat{b}_0$ of parameter b_0. Although the estimate so obtained may be adequate, in principle to do this would be a mistake. To understand why, it is necessary to describe the conceptual framework within which this kind of parameter estimation takes place.

We have denoted m the safety of a specific intersection. Imagine a population of intersections that all have the same traffic flows. In this imaginary population, the m's would still vary from intersection to intersection because, although flows are identical, they involve different drivers in different parts of different cities, and so forth. Thus, one can speak of the mean of the m's ($E\{m\}$) in this imaginary population of intersections with identical traffic flows. This mean of m's is what describes the safety of a "representative" or "average" intersection for this imaginary population. Similarly, one can speak of the variance of the m's (Var$\{m\}$) in this imaginary population of intersections. To make a statement about the safety of a specific intersection of this population, it can be said that "intersections of this kind (e.g., signalized, with a specific pattern of traffic flow) have on the average $E\{m\}$ accidents and the variability of m in the population of intersections of this kind is Var$\{m\}$."

When fitting a model to accident data, we are trying to estimate $E\{m\}$ as a function of traffic flow (in this case). That is, we are trying to determine what the m is of some "average" or "representative" intersection and how it varies with traffic flow. However, the data used for estimation are not for "average" intersections. Each accident count we use is for one specific intersection from the imaginary population of intersections with the same flows. It follows that if $E\{m\}$ is what we wish to estimate, the accident count must be considered as a Poisson random variable that comes from a site with m as its expected value and that this m, in turn, is one of a distribution of m's characterized by $E\{m\}$ and $\text{Var}\{m\}$.

Thus, the distribution of accident counts around $E\{m\}$ is one of a family of "compound Poisson distributions." In the special case in which the distribution of m's in these imaginary populations can be described by a gamma probability density function, the distribution of accident counts around the $E\{m\}$ must be taken as negative binomial. This is a radical departure from the assumptions on which the usual regression software is based. Therefore, a "least-squares" regression model should not be used without expert modification.

In this project we have estimated coefficients using the Generalized Linear Interactive Modeling (GLIM) software package (*17*). This software yields maximum likelihood estimates of coefficients and allows the user to specify the "error structure" that corresponds to the data used. In our case we specified the negative binomial error structure, following Maycock's lead (*18*).

It is already clear that to answer some practical questions about safety it is not sufficient to know what the safety ($E\{m\}$) is of an "average" intersection with given traffic flows. Ordinarily one wishes to make statements about the the safety (m) of some specific intersection. If the $\text{Var}\{m\}$ is very large, knowledge of $E\{m\}$ tells us little about the m of a specific intersection, and vice versa. To make informative statements about the safety of specific intersections, one also needs to know the $\text{Var}\{m\}$. From the methodological point of view, it is the approach to the estimation of $\text{Var}\{m\}$ that is of the most interest.

It can be shown that if accident occurrence follows the Poisson probability law for each intersection in the aforementioned imaginary population, the variance of accident counts in such a population is given by

$$\text{Var}\{\text{accident counts}\} = \text{Var}\{m\} + E\{m\} \quad (2)$$

It follows that one can estimate $\text{Var}\{m\}$ if estimates of $E\{m\}$ and Var{accident counts} are available. An estimate of $E\{m\}$ is the direct product of coefficient estimation. Thus, once we have estimates of b_0, b_1, . . . for a certain accident pattern, we can estimate $E\{m\}$. We explain below how to get an estimate of Var{accident counts}.

Consider one site in our data with its specific traffic flows. The squared difference between the accident count on that site and the corresponding $E\{m\}$ is an estimate of Var{accident counts} for that specific combination of flows. When we plot these squared differences (often called residuals) against $E\{m\}$, we find a relationship of the following form:

$$\hat{\text{Var}}\{\text{accident counts}\} = \hat{E}\{m\} + [\hat{E}\{m\}]^2/k \quad (3)$$

The same relationship was found in the analysis of accidents at grade crossings (*14*) and was suggested by others (*19*) earlier. Thus, taken together, the residuals behave in a sufficiently regular fashion to allow the estimation of the parameter k. Once an estimate of k is available, the estimate sought is given by

$$\hat{\text{Var}}\{m\} = [\hat{E}\{m\}]^2/\hat{k} \quad (4)$$

The process of coefficient estimation that we used is iterative. We begin by assuming a value for k and proceed to estimate the vector of b-coefficients with GLIM. Now we can calculate residuals. These serve as input into a program to obtain the maximum likelihood estimate of k. The new k is fed back into GLIM to obtain new estimates of b-coefficients, and the cycle is repeated until closure.

The most frequent accidents are of Patterns 1, 2, 4, and 6, accounting for 83 percent. Thus it was possible to study these patterns in considerable detail. It was found that the number of accidents varies with the time of day. Therefore, the b-coefficients were estimated separately for the a.m. peak, the p.m. peak, the off-peak, and average daytime (daily) conditions. Although we provide models and coefficient estimates for all other patterns as well, because of the limited data, these should be regarded as unreliable. The final models selected for the 15 accident patterns and their coefficients are given in Table 3.

For clarity, the dimensions of $\hat{E}\{m\}$ (the estimate of $E\{m\}$) are accidents per hour (a/h) and the dimensions of the traffic flows are correspondingly vehicles per hour (v/h). The convenience of this consistency will become evident in the next section.

Not all accidents are "reportable" and not all reportable accidents are in fact reported. During the 1982–1984 period, accidents that exceeded $400 CDN in damages as well as accidents involving injury (visible or complained of) were reportable. These criteria as well as the completeness of reporting vary in time and place. The equations specified by Table 3 allow the estimation of total accidents in Metropolitan Toronto during 1982–1984. To increase the transferability of our findings, we have also provided estimates of the injury ratio (injury accidents/total accidents) for each accident pattern in Table 4.

In Figure 4 we show the ratio of observed accident count per $\hat{E}\{m\}$ for the total number of accidents. The 145 intersections were arranged in ascending order according to $\hat{E}\{m\}$. It is of course pleasing to see the ratios symmetrically distributed around 1. The upper and lower bounds (1.72, 0.37) in Figure 4 contain 90 percent of the ratios. The circles represent the averages of 20 ratios. This is not a great achievement because the same data used earlier for coefficient estimation are now used again to describe model performance.

In Figure 5a and b we show the same ratios (observed/estimated) for accident Patterns 1 and 6. It can be seen clearly that the variability of the ratios decreases as $\hat{E}\{m\}$ increases. The circles represent the averages of 20 ratios. As expected, these averages fall neatly around 1.

Finally, the reader is reminded that the equations in Table 3 are for weekday, daytime accidents at signalized intersections in which two vehicles collided.

TABLE 3 ACCIDENT PREDICTION MODELS

PATTERN	MODEL FORM	TIME	COEFFICIENT ESTIMATES			$\hat{k}$
			b_0	b_1	b_2	
1	$\hat{E}(m) = b_0 \times F$	AM	0.1655×10^{-6}			2.98
		PM	0.2178×10^{-6}			2.73
		off	0.2164×10^{-6}			3.54
		daily	0.2052×10^{-6}			4.59
2	$\hat{E}(m) = b_0 \times F$	AM	0.0987×10^{-6}			1.49
		PM	0.0933×10^{-6}			0.94
		off	0.1080×10^{-6}			4.15
		daily	0.1014×10^{-6}			1.97
4	$\hat{E}(m) = b_0 \times F_2^{b_2}$	AM	19.020×10^{-6}		0.1536	2.65
		PM	1.4127×10^{-6}		0.6044	2.33
		off	9.7329×10^{-6}		0.3860	3.38
		daily	8.1296×10^{-6}		0.3662	5.51
6	$\hat{E}(m) = b_0 \times F_1 \times F_2^{b_2}$	AM	0.0283×10^{-6}		0.5163	1.39
		PM	0.0940×10^{-6}		0.3091	2.70
		off	0.0718×10^{-6}		0.4127	2.20
		daily	0.0418×10^{-6}		0.4634	2.10
3	$\hat{E}(m) = b_0 \times F_2^{b_2}$	daily	8.6129×10^{-9}		1.0682	use $\hat{k}$ = 1, 2
5	$\hat{E}(m) = b_0 \times F_1^{b_1} \times F_2^{b_2}$	daily	0.3449×10^{-6}	0.1363	0.6013	
7	$\hat{E}(m) = b_0 \times F_1^{b_1} \times F_2^{b_2}$	daily	0.2113×10^{-6}	0.3468	0.4051	
8	$\hat{E}(m) = b_0 \times F_2^{b_2}$	daily	2.6792×10^{-6}		0.2476	
9	$\hat{E}(m) = b_0 \times F^{b_1}$	daily	6.9815×10^{-9}	1.4892		
10	$\hat{E}(m) = b_0 \times F_2^{b_2}$	daily	5.590×10^{-12}		2.7862	
11	$\hat{E}(m) = b_0 \times F_1^{b_1} \times F_2^{b_2}$	daily	1.3012×10^{-9}	1.1432	0.4353	
12	$\hat{E}(m) = b_0 \times F_1^{b_1} \times F_2^{b_2}$	daily	0.0106×10^{-6}	0.6135	0.7858	
13	$\hat{E}(m) = b_0 \times F_1^{b_1} \times F_2^{b_2}$	daily	0.4846×10^{-6}	0.2769	0.4479	
14	$\hat{E}(m) = b_0 \times F_1^{b_1} \times F_2^{b_2}$	daily	1.7741×10^{-9}	1.1121	0.5467	
15	$\hat{E}(m) = b_0 \times F^{b_1}$	daily	0.5355×10^{-6}	0.4610		

TABLE 4 INJURY RATIO BY PATTERN

PATTERN	MEAN RATIO	95% CONFIDENCE LIMIT
1	0.3885	0.34 - 0.45
2.	0.1929	0.14 - 0.27
4.	0.3256	0.23 - 0.40
6.	0.3169	0.28 - 0.36
other	0.1685	0.13 - 0.22
TOTAL	0.2989	0.28 - 0.33

NUMERICAL EXAMPLE

To illustrate the use of the equations in Table 3, we examine one intersection in detail. The a.m. peak traffic flows for the intersection are shown in Figure 6. To find the $\hat{E}\{m\}$ for a particular accident pattern, the flows in Figure 6 are substituted in the equations of Table 3. For example, to obtain the $\hat{E}\{m_6\}_{am}$ we use

$$\hat{E}\{m_6\}_{am} = \sum_{j=1}^{4} (0.0283 \times 10^{-6} \times F_{1j} \times F_{2j}^{0.5163})$$
$$= 7.96 \times 10^{-4} \text{ (a/h)}$$

and from Equation 4

$$\hat{\mathrm{Var}}\{m_6\}_{am} = \sum_{j=1}^{4} [\hat{E}_j\{m_6\}_{am}]^2/\hat{k}_{6,am}$$
$$= 1.18 \times 10^{-7} \text{ (a/h)}^2$$

where

$$\begin{aligned} F_1 &= 450 \text{ (v/h)} & F_2 &= 120 \text{ (v/h)} & \text{for } j &= 1 \\ &= 986 \text{ (v/h)} & &= 41 \text{ (v/h)} & &= 2 \\ &= 850 \text{ (v/h)} & &= 96 \text{ (v/h)} & &= 3 \\ &= 869 \text{ (v/h)} & &= 59 \text{ (v/h)} & &= 4 \end{aligned}$$

To find $\hat{E}\{m_6\}$ for 2 a.m. peak hours over 261 weekdays in 3 years, multiply 7.96×10^{-4} by $2 \times 3 \times 261 = 1{,}566$, and multiply the $\hat{\mathrm{Var}}\{m_6\}_{am}$ by $(1{,}566)^2$. The equations in Table 3 are for estimating total accidents; therefore, to determine the number of injury accidents we multiply the estimate by the injury ratio from Table 4. Hence, the

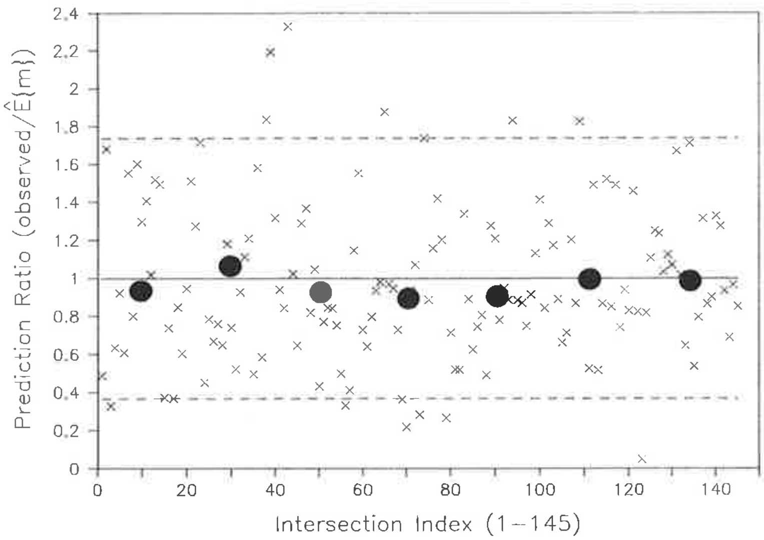

FIGURE 4 Prediction ratio for total number of accidents.

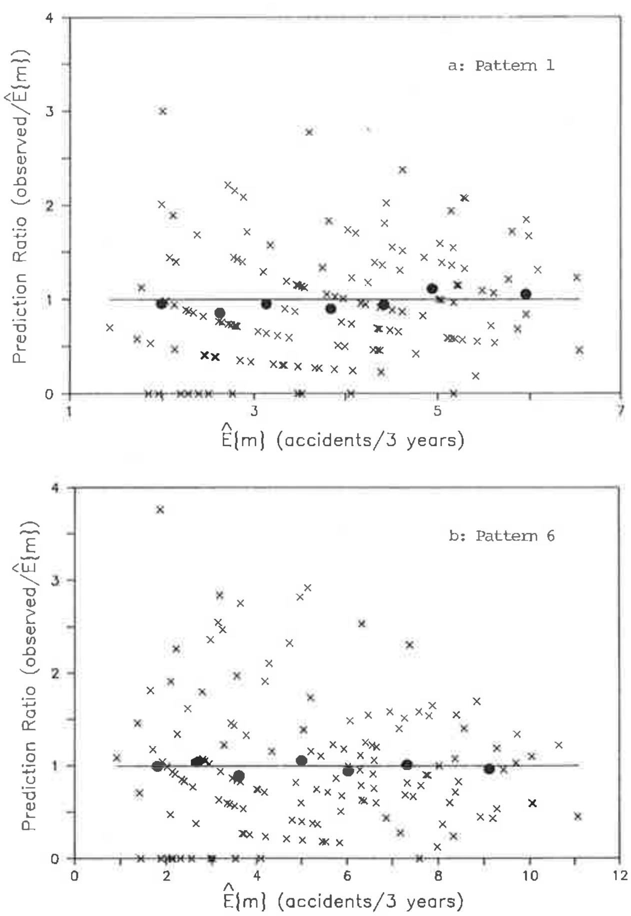

FIGURE 5 Prediction ratio for accident patterns 1 and 6.

estimated mean number of a.m.-peak, Pattern-6 injury accidents for 3 years is

$$\hat{E}\{m_{6,i}\}_{am} = 7.96 \times 10^{-4} \times 1{,}566 \times 0.3169$$
$$= 0.40 \text{ (injury accidents/3 years)}$$

In Table 5 we show the total observed counts, $\hat{E}\{m\}$'s, and $\hat{Var}\{m\}$'s for each accident pattern for 3 years at this intersection. The estimated means were calculated by using the average "daily" models in Table 3 multiplied by a factor of 7,047 (9 × 3 × 261). The reader is reminded that the flows are given in hourly numbers of vehicles. That is,

$$F_{day} = (2 * F_{am} + 2 * F_{pm} + 5 * F_{off})/9$$

The estimated mean and variance for total number of accidents on the intersection are shown on the last row of Table 5. It is the sum of the individual $\hat{E}\{m\}$'s and $\hat{Var}\{m\}$'s for each accident pattern.

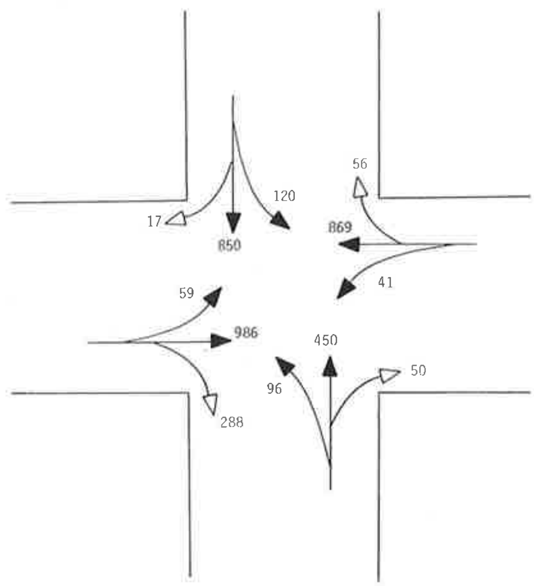

FIGURE 6 Hourly flows for a.m. peak.

ESTIMATION OF SAFETY WHEN ACCIDENT RECORD IS GIVEN

If someone provides us with the traffic flows on a signalized intersection and asks about its safety, we first calculate $\hat{E}\{m\}$ and $\hat{Var}\{m\}$ and then proceed to state: "Intersections

TABLE 5 CORRESPONDENCE BETWEEN OBSERVED AND ESTIMATED NUMBER OF ACCIDENTS

PATTERN	OBSERVED (1982-1984)	$\hat{E}\{m\}$ (3 years)	$\hat{VAR}(m)$ (3 years)
1	4	4.837	1.295
2	2	1.886	0.457
3	0	0.119	0.004
4	1	2.380	0.257
5	0	0.318	0.013
6	6	6.565	5.245
7	0	0.327	0.139
8	1	0.220	0.006
9	0	0.137	0.003
10	0	0.011	0.000
11	0	0.034	0.000
12	1	0.141	0.003
13	1	0.601	0.051
14	0	0.772	0.088
15	0	0.116	0.002
TOTAL	16	18.464	7.563

with such flows are estimated to have, on the average, $\hat{E}\{m\}$ accidents and the variability of m's among similar intersections with such flows is estimated to be $\hat{\text{Var}}\{m\}$."

How does the statement change if, in addition to information about the traffic flows, we are also told that in the last n units of time, X accidents (of a certain pattern) have been recorded?

Before the question is answered, it is instructive to see that the conceptual framework specified earlier still continues to serve. In the section headed Estimation of Coefficients, we have imagined a subpopulation of all intersections such that all its members have the same pattern of traffic flow. This allowed us to make statements about the safety of intersections with a specific flow. The statement in the opening paragraph of this section is about one member of such a subpopulation. Now we are given more information. Not only do we know the flow of traffic for the intersection of interest, we also know its accident record. Thus, the earlier subpopulation (intersections with a certain traffic flow) can be further subdivided in our imagination into still more specific subsets; those with 0 accidents in n units of time, those with 1 accident in n units of time, and so on. The more information we have, the more specific is the subpopulation. For notational clarity we will use $E\{m \mid X,n\}$ and $\text{Var}\{m \mid X,n\}$ to denote the mean and variance of the m's in the subpopulation of intersections with the given flow of traffic and a record of X accidents in the past n units of time.

It can be shown that if the distribution of m's in these imaginary populations can be described by a gamma probability density function (and accident occurrence on any entity follows the Poisson probability law), then

$$E\{m \mid X,n\} = (X + b)/(n + a) \quad (5)$$

$$\text{Var}\{m \mid X,n\} = (X + b)/(n + a)^2 \quad (6)$$

in which,

$$a = E\{m\}/\text{Var}\{m\} \quad (7)$$

$$b = (E\{m\})^2/\text{Var}\{m\} \quad (8)$$

To illustrate this we continue with the earlier numerical example. We found that for the flows given in Figure 6, $\hat{E}\{m_6\} = 7.96 \times 10^{-4}$(a/h) and $\hat{\text{Var}}\{m_6\} = 1.18 \times 10^{-7}$ (a/h)^2. What can be said about the safety of one such intersection on which two weekday, a.m.-peak, Pattern 6 accidents occurred in the last 3 years?

Using the estimates of $E\{m_6\}$ and $\text{Var}\{m_6\}$ above, we find that $\hat{a} = 6{,}745.76$ and $\hat{b} = 5.37$. Therefore, from Equations 5 and 6 we get

$$\hat{E}\{m \mid 2{,}1566\} = 8.87 \times 10^{-4} \quad \text{(a/h)}$$

and

$$\text{Var}\{m \mid 2{,}1566\} = 1.07 \times 10^{-7} \quad \text{(a/h)}^2$$

This can be made more tangible by visual representation. The gamma probability density function is given by

$$f(m) = a^b\, m^{b-1}\, e^{-am}/\Gamma(b) \quad \text{for } m > 0 \text{ and } 0 \text{ otherwise} \quad (9)$$

Curve A in Figure 7 shows the $f(m_6)$ for intersections with flows given in Figure 6. The mean of m's for such intersections is 7.96×10^{-4} (a/h), and the variance is 1.18×10^{-7} (a/h)^2. By using $f(m)$ one can even say, for example, that 5 percent of intersections of this kind have m's in excess of 0.00146 (a/h).

Curve B in Figure 7 shows the $f(m_6 \mid X,n)$ for the same F_1's and F_2's when $X = 2$ accidents and $n = 1{,}566$ hr. To plot this curve we use Equation 9 again, except that a is replaced by $(n + a)$ and b is replaced by $(X + b)$. The mean of m's for intersections carrying these flows and with such an accident record is 8.87×10^{-4} (a/h), and the variance of m's is 1.07×10^{-7} (a/h)^2.

SUMMARY AND DISCUSSION

We have used accident data and information about intersection traffic flows to build models for the estimation of safety at signalized intersections.

During the course of model development we reached some useful insights. First, for a model to portray a relationship between cause (traffic flow) and effect (collisions between vehicles) we chose to relate accidents to the traffic flows to which the colliding vehicles belong. In our view, the logic of attempts to seek an aggregate relationship between accident frequency and some function of all flows (sum of entering flows, sum of products of flows, etc.) is unsatisfactory.

Second, it appears that the customary categorization of accidents by initial impact (rear end, angle, turning movement, sideswipe, etc.) is not very informative. One cannot assume, for example, that classification of an accident as an angle accident implies that the vehicles were traveling at right angles to each other or that most accidents involving left- or right-turning vehicles will be classified as turning accidents.

Third, a close examination of how the frequency of collisions depends on the traffic flows from which they arise reveals that preconceived notions are at times not borne out by empirical evidence. We find that the frequency of collisions between vehicles traveling in the same

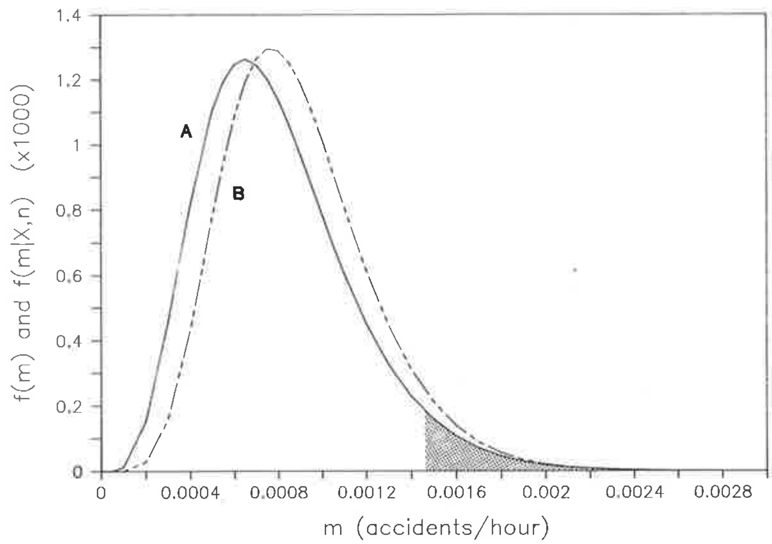

FIGURE 7 Gamma probability density function for *m*.

direction is proportional to the traffic flow in that direction, as one would expect. However, we also observe that the frequency of collisions between vehicles turning left and those proceeding straight through is proportional to the flow of through traffic but less than proportional to the flow of left-turning vehicles. In addition, the frequency of collisions between vehicles traveling at right angles to each other (~12 percent of all collisions) does not seem to depend at all on the larger of the two traffic flows; it increases with the smaller flow, but less than linearly.

Taken together, these observations lead to the conclusion that the popular assumption that intersection accidents are proportional to the sum of entering volumes is not in line with empirical evidence for several common accident types. Therefore, it cannot be true for the totality of intersection accidents. It follows that it is not correct to use intersection accident rates calculated on the basis of the sum of entering volumes to compare the safety of two different intersections nor is it proper to use the sum of entering volumes to correct for exposure in before-and-after studies.

On the basis of such exploratory analysis we have selected plausible model forms that would fit the data with parsimony of coefficients.

The following conceptual frame serves for both coefficient estimation and later for the model use. We think of a specific intersection as being a member of an imaginary population of intersections with similar features. Thus, for example, if we know the accident history and traffic flows for a certain signalized, four-legged intersection in metropolitan Toronto, we imagine a population of such intersections with the same flows and accident history. When making a statement about the safety, m, of a specific intersection, we say that intersections with similar features have on the average $E\{m\}$ accidents and that in this imaginary population, m's have a variance given by $\mathrm{Var}\{m\}$. Thus, the estimate of the population average $[E\{m\}]$ is our best estimate of the m for the specific intersection about which we speak and the estimate of $\mathrm{Var}\{m\}$ describes the uncertainty surrounding this estimate of m.

This sounds like academic hairsplitting and an unnecessary stretching of the imagination. However, the consequences of adopting this approach are immediate, practical, and far-reaching.

First, this approach implies that the run-of-the-mill least-squares regression software should not be used in the analysis of accident data. The error structure must be taken to be the compound Poisson kind. Second, the same approach allows one to use the residuals in order to obtain an estimate of $\mathrm{Var}\{m\}$. The knowledge of $\mathrm{Var}\{m\}$ is essential, as is shown; it allows us to combine data about accidents and information about traffic flows and use both for the estimation of safety. Knowledge of $\mathrm{Var}\{m\}$ is also the basis of all statements about the effect of any safety treatment or about what is "normal" or "unduly hazardous."

The results of our modeling and coefficient estimation efforts are summarized in Table 3. Here we give equations for each of 15 accident patterns to estimate $E\{m\}$ and $\mathrm{Var}\{m\}$ as a function of intersection flows. The use of these equations is demonstrated by a numerical example. We show that given the details of traffic flow at an intersection, we are now in a position to estimate the number of accidents expected to occur per unit of time in each of 15 accident patterns and also to describe the uncertainty surrounding this estimate.

Thus far the estimates have been "personalized" to account for the specific traffic flows at an intersection. The next step is to harness for estimation also the information about the accident history of a specific intersection. The conceptual framework set up earlier again stands in good stead. Simple equations allow the transition from the estimate of $E\{m\}$ and $\mathrm{Var}\{m\}$ based on traffic flows to the corresponding estimates, which are now based on the accident history as well. The underlying assumption is that the distribution of the m's in each imaginary population of intersections with similar features can be described by a gamma probability density function.

Relying on the same assumption, one can now obtain a complete description of the probability density of m's. We have shown this in another numerical example. Thus, we are now in a position to make several informative statements.

First, given the flows of a signalized intersection, we can say how many accidents and what type should be expected to occur on it. We can also specify the variance to be expected; in fact, the complete probability density function can be specified and plotted. On the basis of the probability density function we can decide what an unusually high m would be for accidents of a certain type on intersections with such flows.

Second, we can calculate how the number of accidents by type is expected to change when flows change. This allows us to separate changes in safety due to changes in traffic flow from changes due to other reasons. It also allows the correct cross-sectional comparison of the safety of several intersections instead of incorrect comparisons couched in terms of accidents per million entering vehicles.

Third, we can obtain an estimate of m for an intersection with a known accident history ($E\{m \mid X,n\}$) and compare it with what is expected at an average intersection with such flows ($E\{m\}$). It is on the basis of the magnitude $E\{m \mid X,n\}$ and its comparison with $E\{m\}$ (and also the two corresponding variances) that one should make decisions on what is deviant and where remedial action is warranted.

In particular, some warrants in the Manual on Uniform Traffic Control Devices (e.g., 2B-6 and 4C-8) refer to an "accident problem" and suggest that this is indicated by a certain "number of reported accidents of a type susceptible to correction" that occur in a 12-month period. In our view, such warrants should refer to estimates of $E\{m\}$ and $E\{m \mid X,n\}$ and their comparison and not to the count of reported accidents in a relatively short period of time.

ACKNOWLEDGMENT

Work on this subject has been made possible by the support of the National Science and Engineering Research Council of Canada, Transport Canada (Road Safety), and the Federal Highway Administration, U.S. Department of Transportation. The data for the analysis were provided by the Metropolitan Toronto Roads and Traffic Department.

REFERENCES

1. R. A. Chapman. The Concept of Exposure. *Accident Analysis and Prevention*, Vol. 5, 1972, pp. 95–110.
2. S. P. Satterthwaite. *A Survey of Research into Relationships Between Traffic Accidents and Traffic Volumes*. TRRL SP 692. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1981.
3. J. D. Thorpe. Traffic Accident Measurements. *Journal of the Institution of Engineers*, Australia, Vol. 35, 1963.
4. W. L. Smith. Probability Study of High-Accident Locations in Kansas City, Missouri. *Traffic Engineering*, Vol. 40, No. 7, 1970, pp. 42–49.
5. G. Worsey. Predicting Urban Accident Rates from Road and Traffic Characteristics. *ITE Journal*, Dec. 1985.
6. S. M. Breunning and A. J. Bone. Interchange Accident Exposure. In *Highway Research Board Bulletin 240*, HRB, National Research Council, Washington, D.C., 1959, pp. 44–52.
7. V. H. Surti. Accident Exposure for At-Grade Intersections. *Traffic Engineering*, Vol. 36, No. 3, 1965, pp. 26–27, 53.
8. A. S. Hakkert and D. Mahalel. Estimating the Number of Accidents at Intersections from a Knowledge of the Traffic Flows on the Approaches. *Accident Analysis and Prevention*, Vol. 10, 1978, pp. 69–79.
9. J. W. McDonald. Relationship Between Number of Accidents and Traffic Volumes at Divided Highway Intersections. In *Highway Research Board Bulletin 74*, HRB, National Research Council, Washington, D.C., 1966, pp. 7–17.
10. G. M. Webb. The Relationship Between Accidents and Traffic Volumes at Signalized Intersections. *Proc., Institute of Traffic Engineers*, 1955, pp. 149–160.
11. J. C. Tanner. Accidents at Rural Three Way Junctions. *Journal of the Institution of Highway Engineers*, Vol. 2, No. 11, 1953, pp. 56–67.
12. J. Coleman and G. R. Stewart. Investigation of Accident Data for Railroad-Highway Grade Crossings. In *Transportation Research Record 611*, TRB, National Research Council, Washington, D.C., 1976, pp. 60–67.
13. P. Mengert. *Rail-Highway Crossing Resource Allocation Procedure: User's Guide*, 2nd ed. Report FHWA-IP-86-11. FHWA, U.S. Department of Transportation, July 1986.
14. E. Hauer and B. N. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effect of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1988, pp. 131–140.
15. C. V. Zegeer, J. G. Mayes, and R. C. Deen. *Cost-Effectiveness of Lane and Shoulder Widening of Rural, Two-Lane Roads in Kentucky*. Bureau of Highways, Kentucky Department of Transportation, July 1979 (also published in *Transportation Research Record 806*, TRB, National Research Council, Washington, D.C., 1981).
16. E. Hauer, E. Pagitsas, and B.-Y. T. Shin. Estimation of Turning Flows from Automatic Counts. In *Transportation Research Record 795*, TRB, National Research Council, Washington, D.C., 1981, pp. 1–7.
17. R. J. Baker and J. A. Nelder. *The GLIM System: Release 3, Manual*. Rothhamsted Experimental Station, Harpenden, England, 1978.
18. G. Maycock and R. D. Hall. *Accidents at 4-Arm Roundabouts*. TRRL Laboratory Report 1120. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1984.
19. T. P. Hutchinson and A. J. Mayne. The Year-to-Year Variability in the Number of Road Accidents. *Traffic Engineering and Control*, Vol. 18, No. 9, Sept. 1977.

DISCUSSION

CARL N. MORRIS
Department of Mathematics and Statistical Sciences Center, University of Texas, Austin, Tex. 78712.

The idea by Hauer et al. of estimating traffic accident rates based on relevant flow rates certainly must be a welcome advance. But I will leave further assessment of that topic to others and concentrate on the statistical issues, about which I am most familiar.

The authors build on their weighty past contributions to accident analysis by again developing hierarchical and empirical Bayesian models for accident rates, this time further including regression model features. I will summarize their statistical model within a very useful paradigm, which I sometimes call the General Model for Statistics (*1, 2*) because it contains the main branches of statistical thought. Without going through the general ideas, I immediately specialize that model to the Poisson-gamma model of Hauer et al., using their notation. The model has two completely equivalent renderings, a descriptive model that describes how we visualize distributions for the data and parameters, and an inferential model that rewrites the distributions in more convenient form for making inferences. Table 6 will help to clarify my questions.

The descriptive model is mathematically equivalent to the inferential model (Table 7), which specifies the marginal distribution for the data and the conditional distribution for the parameters.

Equation 14 in Table 7 declares the marginal distribution of X to be a negative binomial distribution, with usual parameters $b > 0$ and $p = n/(n + a)$. Expression 15 specifies the mean and variance of this negative binomial distribution inside the square brackets. Similarly, the square-bracket notation in Equation 18 indicates the mean and variance.

I now ask several questions that arise from consideration of the authors' methodology as structured in Equations 10–18.

TABLE 6 DESCRIPTIVE MODEL

Distribution for Observed Data:	$X \mid m \sim \text{Poisson}(mn)$	(10)
	X = accidents at a site.	
	m = true accident rate.	
	n = exposure (time).	
Distribution for Unobserved Parameters:	$m \mid b, a \sim \text{Gam}\left(b, \frac{1}{a}\right)$	(11)
	$Em = b/a = b_0 F_1^{b_1} F_2^{b_2} = \mu$,	(12)
	$\text{Var}(m) = b/a^2$.	(13)
	ϕ = unknown hyperparameters = (b, b_0, b_1, b_2) .	

TABLE 7 INFERENTIAL MODEL

Distribution for Observed Data:	$X \sim \text{NegBin}\left(b, p \equiv \frac{n}{n+a}\right)$	(14)
	$= \text{NegBin}\left[\frac{nb}{a}, \frac{n^2 b}{a^2}\right] = \text{NegBin}\left[\mu, \mu + \frac{\mu^2}{b}\right]$	(15)
	$\mu = b/a = b_0 F_1^{b_1} F_2^{b_2}$.	(16)
Distribution for Unobserved Parameters:	$m \mid X \sim \text{Gamma}\left(b + X, \frac{1}{n+a}\right)$	(17)
	$= \text{Gamma}\left[\frac{b+X}{n+a}, \frac{b+X}{(n+a)^2}\right]$	(18)

1. Parameters b and k: The authors have an expression similar to my Expression 15 for $\text{Var}(X) = \mu + \mu^2/b$, which is their Equation 3 of the section headed Estimation of Coefficients. But they use k for b. Isn't $k = b$? In Pattern 6 they estimate $k = 1.39$ in the morning (Table 3) but $b = 5.4$ (Figure 7) for the same data. Shouldn't these values agree?

2. The link function: The regression coefficients b_0, b_1, b_2 are estimated by the authors using GLIM, which requires, among other things, specification of the so-called "link function" $\eta = g(\mu)$. Here η, the linear form of the model, is

$$\eta = \log(b_0) + b_1 \log(F_1) + b_2 \log(F_2) \tag{19}$$

The mean structure therefore satisfies

$$\mu = EX = \frac{nb}{a} = b_0 F_1^{b_1} F_2^{b_2} = \exp(\eta) \tag{20}$$

and so, presumably, we are required to use

$$\eta = \log(\mu) \tag{21}$$

which is the "log link." Is this correct? Note that the "natural link" for the negative binomial family,

$$\eta = g(\mu) = \log\left(\frac{\mu}{b + \mu}\right) = \log\left(\frac{n}{n + a}\right) \tag{22}$$

would be fit by GLIM if no specification were made. But the natural link is not the link in Equation 21.

3. Variable exposures (n): Do the exposures n in Equation 10 vary from intersection to intersection? Assuming so, should that complicate the estimation of b and of the coefficients (b_0, b_1, b_2) through Equation 13?

4. Details about GLIM: Models like Equation 14 are quite difficult to fit. The authors seem to have found an ingenious way to estimate the "hyperparameters" (b, b_0, b_1, b_2) in the model. (Note that a can be defined in terms of the other paramters by using Equation 16.) More details, beyond the material surrounding the authors' Equation 4, are needed for an adequate understanding of this procedure. Most particularly, the subscripts that indicate the specifics of the intersection are not shown by Hauer et al., and so are also avoided in my rendering above. The clarification would reveal which values are intersection dependent and specify the assumed probabilistic independence of the data $\{X_i\}$ and of the parameters $\{m_i\}$.

5. Need for dependence of the mean on both F_1 and F_2: Although the points lie on a horizontal line in Figure 3a, the fitted curve must go through the origin (no flow implies no accidents). Thus, forms like $b_0 F_2 b_2$ used for Pattern 4 do not seem reasonable because they predict a substantial number of accidents when $F_1 = 0$.

6. Errors in variables: Substantial errors must be made in estimating (F_1, F_2). If so, this produced an errors-in-variables bias in estimation of b_1, b_2. Is there any way to determine how large this bias might be, or how accurately the flows are measured?

In conclusion, the paper offers a potentially very useful methodology that should further improve assessment of traffic hazards. The answers to the foregoing questions will help me to further understand the approach offered and its advantages.

ACKNOWLEDGMENT

This discussion was based on research supported by FHWA at the Texas Transportation Institute, Texas A&M University, College Station, Texas. Olga Pendleton is the principal investigator.

REFERENCES

1. C. N. Morris. Parametric Empirical Bayes Inference: Theory and Applications (with Discussion). *Journal of the American Statistical Association*, Vol. 78, No. 3, 1983, pp. 47–65.
2. C. N. Morris. Determining the Accuracy of Bayesian Empirical Bayes Estimators in the Familiar Exponential Families. In *Statistical Decision Theory and Related Topics*, IV, Vol. 1 (S. S. Gupta and J. O. Berger, eds.), Springer-Verlag, New York, 1988, pp. 251–263.

The opinions are those of the author alone, and have not been reviewed by FHWA.

AUTHORS' CLOSURE

It is most gratifying to have a prominent statistician review and comment on our work. We will attempt to respond to the questions raised.

PARAMETERS *b* AND *k*

There is no natural reason why in some population of intersections the *m*'s should fit a gamma distribution. However, if they do, the plot of squared residuals should be of the form specified by Equation 3. Having reassured ourselves that using the gamma assumption is in accord with our data, we estimated the value of k. Indeed, the estimate of k is also the estimate of b, as pointed out by the discussant.

As pointed out by the discussant also, for Pattern 6 (a.m. peak), $\hat{k} = 139$, whereas a few pages later, for the same pattern, we use $\hat{b} = 5.37$, an apparent discrepancy.

To see why the two values must differ and how, consider the characteristic function of a gamma distribution,

$$\phi(t) = 1/(1 - it/a)^b$$

If four independent random variables have the same gamma distribution, their sum will have a characteristic function,

$$\phi^*(t) = 1/(1 - it/a)^{4b}$$

Indeed, the 1.39 ($\hat{k}$) value in Table 3 refers to one gamma-distributed variable. The 5.37 ($\hat{b}$) used later pertains to a sum of four random variables (similar but not identical). Thus the usage does seem to be correct.

QUESTIONS ABOUT GLIM

The GLIM software has been used for this purpose by others before us [e.g., Baker and Nelder (*1*), Pickering et al. (*2*), and Hall (*3*)]. That is why we decided not to "attempt a comprehensive coverage of technical detail" in our paper (section on Estimation of Coefficients). Because the frame of reference has not been provided in the paper, it is difficult to give an intelligible discussion here.

In answer to Questions 3 and 4, accident count, duration of accident history, and traffic flow data are all intersection specific and are so represented for use in GLIM.

We have elected to measure μ in accidents per hour and F in vehicles per hour. Suppose now that we have at a certain intersection 3 years' worth of accident data for two morning peak hours during 250 weekdays a year. The n for this intersection would be $3 \times 2 \times 250$. For an intersection i,

$$n_i\mu_i = n_i \, \Pi \, F_j^{b_j}$$
$$\eta_i = \ln(n_i) + \sum [b_j \times \ln(F_j)]$$

In GLIM, $\ln(n_i)$ is treated as an "offset."

NEED FOR F_1 AND F_2

The discussant notes that if $b_0F_2^{b_2}$ is used, one would predict $\mu > 0$ even when $F_1 = 0$. Recall that in the section of our paper on how accidents depend on contributory traffic flows, we called the larger flow F_1 and the lesser flow, F_2. Therefore, $F_1 = 0$ implies that $F_2 = 0$ and "no flow implies no accidents," as is required.

ERRORS IN VARIABLES

As noted by the discussant, F_1 and F_2, which are taken to apply to a 3-year period, are actually estimated from a 1-day volume count. Thus, although GLIM and other statistical software treat the independent variables as if they were measured accurately, in fact they are also subject to error. Some ramifications of the "error-in-variables" problem have been explored recently by Weed and Barros (*4*). Unfortunately, we still do not know what effect this might have on estimates of b_1 and b_2.

REFERENCES

1. R. J. Baker and J. A. Nelder. *The GLIM System: Release 3, Manual.* Rothhamsted Experimental Station, Harpenden, England, 1978.

2. D. Pickering, R. D. Hall, and M. Grimmer. *Accidents at Rural T-Junctions*. TRRL Research Report 65. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1986.
3. R. D. Hall. *Accidents at Four-Arm Single Carriageway Urban Traffic Signals*. TRRL Contractor Report 65. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1986.
4. R. M. Weed and R. T. Barros. Demonstration of Regression Analysis with Error in the Independent Variable. In *Transportation Research Record 1111*, TRB, National Research Council, Washington, D.C., 1987, pp. 48–54.

Environmental Quality and Safety Assessment of Residential Streets

A. S. HAKKERT AND A. H. PISTINER

The use of behaviorally oriented criteria to assess environmental quality of residential streets is considered in this study. The two main aims of the study were (*a*) to establish objectives for environmental design based on residents' perceptions of traffic and design variables and (*b*) to develop a subjective procedure for assessing street environmental quality and to verify the extent to which it agrees with existing procedures developed in Europe. Ten one-way residential streets in Haifa were selected for detailed study. The method of study drew on two sources of information. First, 147 residents of the 10 selected streets responded to a questionnaire in which perception levels regarding a number of street attributes were assessed. Second, a set of systematic observations and objective measurements of traffic and environmental variables was conducted for each street. Multivariate analysis of interview responses, traffic-related parameters, and environmental indicators was conducted to understand better the way in which factors tend to cluster and to develop quantitative relationships from regression analyses of responses to various conditions. The questionnaire data were summarized using the principal components model of multivariate analysis. Four composite variables were extracted and defined from the original 11 variables. Stepwise multiple regression analysis with environmental and traffic variables as predictors yielded correlations of the order of 0.60, indicating that individual perception can be satisfactorily predicted. Street rankings on the basis of environmental quality were assessed according to three different procedures. A fair degree of agreement was achieved (R = 0.70).

The street is multifaceted, at once a channel for traffic and a social space. This is particularly true of residential streets, which are the setting of a large number of human activities, both indoors and out.

By means of an effective design, the street can provide an environment that supports human behavior and supplies its needs. It may be assumed that an environment that provides conditions for the proper performance of dwellers' indoor activities as well as opportunities for human contact and outdoor activities is not only desirable but necessary for environmental (residential) satisfaction.

Street design may affect the quality of the residential environment in two principal ways. First, the street pattern and geometric characteristics will influence, to a large extent, the way in which motor traffic uses the street. Under certain circumstances, motor traffic, both moving and parked, may create inhibiting effects on activities such as walking, sitting, and playing.

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

Indoor activities may also be disturbed by traffic noise and fumes, and a sense of intrusion on privacy may occur as the result of perceived high density (crowding) caused by sensory stimuli such as high traffic density and much parking (*1*). Over the years, a wide range of techniques and design solutions has been developed to restrict traffic volumes and speed and to control parking so that noise, fumes, visual intrusion, pedestrian-vehicle conflicts, and other unpleasant stimuli are reduced to tolerable or desired levels.

Second, design will obviously influence the street's visual appearance, its shape and proportions. Human perception includes the aesthetic experience, which is of ultimate emotional significance. Many streets are flat, uninteresting, and lacking in individual identity. They may lack an appropriate sense of enclosure, stretching visually to infinity through open vistas.

In order to determine the physical dimensions that will give the street a harmonious scale, it is important to consider the dynamic quality of the street (a space that is apprehended), the human scale, and the expected intensity and type of activities that will occur there.

Safety is an important aspect of street life and is generally expressed in terms of accidents, but these are very sparse on any one residential street and are difficult to relate to scientifically or treat statistically. For this reason, surrogate objective variables, such as speed, are adopted. It is, however, important to determine how these surrogate variables are related to safety and also how they are related to other street qualities important to residents.

The general aim of this paper is to explore the relationship between residents' perceptions of street quality and the objective variables related to the various aspects of a street's environmental quality.

The way in which people use the street and for what purposes, as well as the way that residents feel about their home surroundings, will depend a great deal on the manner in which the residential environment is perceived and the relative importance placed upon the factors involved by the people who are subjected to them. Research on how the perception of the impact of roads and traffic is structured by those who are exposed to it suggests that there are three main aspects perceived: severance aspects, including difficulty and danger in crossing the road; noise and fumes (unpleasant conditions); and visual effects (*2*).

Interference with sleep and speech are frequently cited as major reasons for annoyance with residential traffic noise. Robinson (*3*) states that a sound level of 48 dB(A) allows conversation in a normal voice at a distance of 4 m. Beranek (*4*) quotes a maximum indoor level of 40 to 45 dB(A) if television and radio are to be understood comfortably. Subjective ratings of motor vehicle noise heard outdoors are provided by Wilson (*5*). Sound levels below 70 dB(A) are considered quiet, whereas levels between 80 and 88 dB(A) are rated noisy, and levels above 89 dB(A) are rated very noisy. Further details on the relationship between traffic characteristics and noise are provided by Alexandre et al. (*6*), the U.K. Department of the Environment (*7*), and Hothershall and Slater (*8*). Traffic fumes are less noticeable to those indoors but are found to be a great nuisance to those outdoors. Harmful effects generally attributed to traffic fumes include chest pain, poisoning of the atmosphere, breathing afflictions, cancer, sore throats, and coughs (*2*).

The effects of roads and traffic on vision are more complex than those on hearing or smell. However, some of the work carried out in the field, which combined physical measures of visual intrusion with social surveys, shows that factors such as the amount of road, traffic, sky, and greenery visible from the observer's angle of view are significant in determining visual satisfaction (*9*).

The danger of traffic is generally measured by the occurrence of accidents. However, in residential areas, accidents are a rather unsatisfactory indicator for lack of traffic safety. A further dimension in the problem of pedestrian safety on residential streets is the level of safety perceived by those who live on that street. Casualties being relatively rare events, residents do not generally think of their streets as being unsafe in terms of frequent collisions; however, they express a broadly perceived concern about "traffic problems," almost always related to traffic speed and volume, which are closely related to accident potential (*10*). In a recent study on subjective safety conducted in Holland (*11*), which included both questionnaire surveys of residents and objective measurements, it was found that subjective safety as expressed by residents was most strongly related to traffic volume and speed.

Considering the subjective aspects of all these effects, it might be said that in order to conduct a proper evaluation of street problems, whether related to the environment or traffic conditions, it is essential to supplement the professional evaluation with evaluations by residents.

Quantitative measures must be developed on the basis of social surveys and situational characteristics so that an efficient tool may be provided for the determination of design measures or other policies concerning residential environment improvements. Some attempts in this direction have been made by different researchers and will be mentioned later in this paper.

STUDY OBJECTIVES AND METHOD

This study is concerned with using behaviorally oriented criteria to assess environmental quality and with identifying design variables that have the potential to modify the impact of traffic, not only by actually controlling it, but also by affecting residents' psychological feelings or perceptions. The main aims of the study are as follows:

- To examine a number of subjective criteria suggested for the evaluation of environmental quality;
- To explore the relationship between subjective variables representing residents' perception of street qualities and measured objective variables;
- To develop a model of street environmental quality based on residents' perceptions; and
- To examine the relationship between the model and comparable methods developed in Europe.

In order to examine the perceived effects of specific design variables, a number of streets were selected for study. Because it was not financially possible to look at a large number of situations, it was decided to control certain conditions in the selection of streets. Ten one-way residential streets in Haifa, Israel, were selected on the basis of the following criteria: mainly residential land use, location within one section of the city, relatively homogeneous population by social class and income, traffic volumes up to 250 vehicles/hr during peak hours, one-way traffic flow, and presence of on-street parking. The major environmental differences among the streets were their length, cross-section dimensions, horizontal and vertical alignment, visual character, landscaping, traffic speed and composition, and parking density.

The 10 streets selected from a larger sample of residential streets in Haifa can be described as follows.

- Land use: mainly residential with mostly multifamily apartment buildings, generally six to eight apartments per building.
- Land value: generally among the higher values in the Haifa metropolitan region; all streets are located within one geographical section of Haifa.
- Population profile: relatively homogeneous by social class and income, with high rates of home and car ownership.

One-way streets were selected because a large number of the streets in the sample were one way, and it was believed advantageous not to complicate the study with one more variable. One-way versus two-way operation was, however, not a major consideration of this study. It is believed that the results will be generally applicable and not restricted to one-way streets.

The method of study drew on two sources of information for each street: an attitudinal survey of residents and a set of systematic observations and objective measurements of traffic and environmental conditions.

Attitudinal Survey

Approximately 15 randomly chosen residents on each street were selected for 147 home interviews lasting about 20 min each. The number of interviews carried out represents about 5 percent of the number of households on the streets. The main issues explored in the interviews were residents' perception of the street as a community (for outdoor activities and social contact between neighbors); general residential satisfaction; perceived noise and fumes from traffic; sense of safety from traffic; perception of traffic conditions; and perception of visual and aesthetic street qualities, walking conditions, and comfort.

The interviews were conducted by means of a well-structured questionnaire that allowed the subjects to make their own ratings.

Perception levels were assessed regarding a total of 11 street attributes: traffic safety, volume, speed, noise, and fumes; street use by children (street play); street use by neighbors (sitting, walking, etc.); contact among neighbors; street visual quality; pedestrian comfort; and general residential quality.

A series of continuous rating scales was used to assess residents' perceptions. The scales represent attitude continua upon which subjects were invited to mark their own ratings.

Three basic relationships were formulated, the first two of which were assumed and the third was hypothesized:

1. The level of a resident's satisfaction with an isolated street attribute is a function of the perceived effect of the attribute measured for that respondent;
2. A resident's overall satisfaction with the street environment is a function of his level of satisfaction with the individual attributes of the street; environmental quality can be assessed in terms of user satisfaction; and
3. Residents' perceptions will be related to objective measures of traffic and street conditions, and satisfaction levels will vary from one site to the other according to different street characteristics.

Data Collection (Street Survey)

A series of objective measures was needed that could be used as a setting within which to interpret residents' subjective responses. The following street environmental and traffic conditions were postulated to have a potential impact on residents' perceptions: street length, horizontal and vertical alignment, cross-section dimensions (sidewalks, green strips, traffic and parking lanes, building setback), traffic volume and speed, bus service, parking density, landscaping, street enclosure, noise levels, and residential density.

The objective variables to be subjected to statistical analysis were obtained partly from street measurements and partly from street plans and wide-angle photographs taken along each street (photogrammetric analysis). Table 1 details the complete list of objective variables considered.

TABLE 1 OBJECTIVE VARIABLES

No.	Description	Unit
X6	Street length	Meters
X7	Horizontal alignment, expressed as route factor *R* = over-the-route distance (m)/"airline" distance (m)	Ratio *R* (dimensionless)
X8	Vertical alignment, expressed as dominant gradient along street (more than 50 percent of length)	Percentage
	Cross-section dimensions	
Y2	Sidewalk width	Meters
Y3	Moving-lane width	Meters
Y5	Ratio of (sidewalks + green-strip width) to (traffic-lane + parking-lane width)	Ratio (dimensionless)
Y6	Peak-hour traffic volume	Vph
Y7	Average daily transit traffic	Buses per day
Y8	Average traffic speed	Km/hr
Z2	Speed not exceeded by 85 percent of drivers	Km/hr
Z3	On-street parking density	No parked cars in 100 m
	Photogrammetric analysis	
Z4	Road and traffic intrusion in visual field	Planimeter tool units (not to scale)
Z5	"Openness" (measure of green space enclosure)	
Z6	Landscaping (amount of greenery in visual field)	
Z7	Ratio *Z4* to *Z6*	
W3	Noise level (L10) not exceeded for 10 percent of time	Decibels (A weighted)
W4	Average building set back from lot line	Meters
W6	Residential unit density	Units per 100 m
W7	Building density	Buildings per 100 m
W8	Distance between opposing buildings	Meters

Variables *Z4* (road and traffic intrusion), *Z5* ("openness"), and *Z6* (landscaping) were measured from photographs taken down the center of each street at 200-m intervals. The size of specific areas on the photograph was measured by planimeter, but could easily have been expressed as a percentage of the total area.

The methods employed for data collection included field observations supplemented with manual and mechanical counts and use of existing data sources.

STATISTICAL ANALYSIS AND RESULTS

Multivariate analysis of interview responses, traffic-related parameters, and environmental indicators was conducted to understand better the way in which factors tend to cluster and to develop quantitative relationships from regression analyses of responses to various conditions.

Subjective Data Analysis: Principal Components Model

The scores obtained from the questionnaire for each subject and scale were subjected to statistical analysis. First, an 11 × 11 sample correlation matrix was obtained and used as a basis for extracting principal components (based on 147 observations). The principal components model accounts for the variance within a data set by providing those linear combinations of correlated variables that maximize the variance of the weighted sum. The new variables (the weighted sums) are orthogonal to each other, and therefore independent. In a subspace of dimension less than that defined by the complete set of correlated variables, the phenomenon of interest may be studied more conveniently.

The foregoing procedure extracted four composite variables from the 11 original ones. These composite variables provide the categories that form the residents' perceptive structure of interrelationships, and the weights associated with the correlated attributes (variables) reveal the importance placed on each attribute. The combinations of variables and their definitions are given in Table 2.

The weighted sum of the first variable set accounted for the greatest amount of variance within the data (24 percent).

From the relationships exhibited in the data structure, a number of conclusions can be drawn:

1. The issues approached in the questionnaire were grouped into four main categories of residents' perceptions: traffic safety, aesthetics and comfort, street community, and traffic nuisance.

2. From the correlations between variables within categories, the following conclusions may be drawn:
   a. The perception of risks is related to the perception of traffic conditions, particularly traffic speed. The correlation between perception of speed and perception of risk (n = 147) was .56 (p < .01). Attribute theory describes, among others, the ways in which people attribute unpleasant phenomena (for example, accidents) to their own behavior, the behavior of others, chance, or circumstances. In view of the foregoing results, traffic conditions appear to be causal attributes that largely amount for residents' perception of risk.
   b. The perception of visual quality is related to the perception of comfort and residential quality. This is in agreement with findings reported by Lynch (*12*):

> Perception includes the aesthetic experience but is also entangled with many other purposes: comfort, human interaction, orientation and the communication of status. . . . The aesthetic experience is conveyed not only by vision but by other senses as well. Sensations of cold and moist, hot, bright, windy, warm or sheltered, sound and smell can hardly be separated from the visual impression.

The correlation between perception of visual quality and pedestrian comfort (n = 147) was .51 (p < .01), and the correlation between perception of residential quality and pedestrian comfort was .41 (p < .01).
   c. The perception of street use for outdoor activities is related to the degree of socialization indicated by residents. This is in agreement with the ideas expressed by Jacobs (*13*) and by Unterman (*14*): "The casual sidewalk life provides the natural setting for the public contact between neighbors; enhances socialization and sense of community." The correlation between perception of street use by children and perception of street use for walking, sitting, talking, and other outdoor activities (n = 147) was .41 (p < .01). The correlation between perception of street use for outdoor activities and the degree of socialization between neighbors indicated by residents (n = 147) was .30 (p < .01).
   d. The perception of traffic emissions is related to the perception of traffic volumes. Although other traffic attributes, for example, speed, also affect traffic emissions, only volume was perceived as a causal attribute. The correlation between the perception of noise and traffic volume was .31 (p < .01).

3. The amount of variance explained by each variable cluster points out the importance of Cluster 1, traffic safety, which accounts for almost 25 percent of the variance exhibited, followed by Cluster 2, aesthetics and comfort, which accounts for 16 percent of the variance.

TABLE 2 VARIABLE COMBINATIONS AND DEFINITIONS

Composite Variable	Variables Combined	Definition
C_1	Traffic speed Traffic volume Traffic danger	Traffic safety
C_2	Visual quality Pedestrian comfort Residential quality	Aesthetics and comfort
C_3	Child street-play Outdoor activity Neighboring	Street community
C_4	Traffic fumes Traffic noise Traffic volume	Traffic nuisance

Sample distributions and cumulative frequencies according to C_1, C_2, C_3, C_4 (within streets and in general) provided a valuable insight into the problem areas from a resident's perspective and revealed that responses varied significantly from street to street. The general situation in the study area can be described as follows:

1. More than 50 percent of those interviewed found safety and traffic conditions on their streets unsatisfactory;
2. Approximately 25 percent of those interviewed were not satisfied with their street's appearance and walking conditions;
3. More than 50 percent of those interviewed indicated that nothing much was going on in their streets and knew few or none of their neighbors;
4. Less than 15 percent of those interviewed complained about noise and fumes from traffic on their streets.

Clearly, there is wide concern with traffic conditions among the residents interviewed. Speed, dangerous curves, and lack of visibility were the main causes mentioned for lack of safety, followed by traffic volumes.

In relation to street aesthetics and pedestrian comfort, the main complaints were lack of vegetation, lack of continuity in the walking path, parked cars, large asphalt areas, and microclimate (too hot or too windy).

Development of Qualitative and Quantitative Relationships

The relationships between situational characteristics and perception of street conditions were considered on two levels. First, the individual perception was used as a starting point ($n = 147$ individuals). Second, the average perception per street ($n = 10$ streets) was used as a measure for assessing street environmental quality in terms of user satisfaction.

It would seem important to explore the relationships between residents' subjective assessments, as expressed by the composite scores developed, and the objective variables measured. Through an understanding of these relationships, ways can be found to improve residents' subjective satisfaction with the environmental quality.

Multiple regression analyses were used to determine the main variables that affect residents' perceptions and to determine their relative importance.

Because the multiple regression procedure requires the predictors to be independent, the objective variables submitted for analysis were arranged into subsets so that variables within each subset were independent. Each subset of variables was alternatively subjected to analysis using the SPSSX statistical package; the equations that best explained the variance of the subjective variables are as follows (numbers in parentheses indicate significance levels; levels below 1 in 10 are considered significant):

$$C_1 = \underset{(.029)}{-53.5X_7} - \underset{(.000)}{3.5Z_2} - \underset{(.028)}{8.2X_8} + \underset{(.000)}{417.4}$$

$$R = 0.62 \qquad n = 147$$

where

C_1 = individual perception of traffic and safety,
X_7 = horizontal alignment (curvature) ($2 < X_7 < 3$),
X_8 = vertical alignment (steepness) (2 percent $< X_8 <$ 7 percent),
Z_2 = 85th-percentile speed (35 km/hr $< Z_2 <$ 52 km/hr), and
constant = 417.4.

$$C_2 = -\underset{(.084)}{7.4Y_3} - \underset{(.007)}{.11Z_5} - \underset{(.006)}{159.9Z_3} + \underset{(.000)}{304} \quad R = 0.63$$

where

C_2 = individual perception of aesthetics and comfort,
Y_3 = effective traffic lane (3.4 m $< Y <$ 7.8 m),
Z_5 = "openness" (measure of street enclosure obtained by photogrammetric analysis) ($180 < Z < 505$ planimeter tool units),
Z_3 = on-street parking density ($0.24 < Z < 0.53$ cars/m of street), and
constant = 304.

Stepwise regression analysis over all respondents ($n = 147$) with the measured variables of street environment as predictors yielded a multiple correlation (R) of .62 with the variable perception of traffic safety, C_1 (C_1 being the statistical combination representing the three subjective variables traffic speed, traffic volume, and traffic danger).

The independent variables included in the model were street curvature, street steepness (gradient), and traffic speed. They contribute negatively to the perception of traffic safety.

The multiple correlation (R) with the variable perception of aesthetics and comfort (C_2) was .63. The independent variables included in the model were effective traffic lane width, "openness" (a measure of the street enclosure), and on-street parking density. They contribute negatively to the perception of street aesthetics and comfort (see regression equation for C_2 above), C_2 being the statistical combination representing the three subjective variables visual quality, pedestrian comfort, and residential quality.

The regression analysis did not yield significant correlations with the variables perception of street community C_3) and perception of traffic nuisance (C_4). Individual perceptions of these two variables could not be predicted satisfactorily on the basis of the environmental variables within the range of study. One possible explanation for this result is that all streets selected had reasonably pleasant conditions with regard to noise and fumes, which therefore were not perceived as problems.

Effect of Environmental Conditions on Traffic Behavior

Assuming that environmental conditions have an effect on traffic behavior, the correlations between the objective variables were obtained. Their examination focused on the effect of variables such as traffic lane width, street length, parking density, and horizontal alignment on traffic speeds. The results are discussed in the following paragraphs.

The correlation obtained between traffic speed and traffic lane width supports the theory that a street with wide lanes invites faster movements. The correlation with the 85th-percentile speeds, a measure that accounts for the faster traffic, was slightly higher than that with the mean speed ($r = .72$, $p < .01$). Effective lanes 3.5 to 4.5 m wide were found to have midblock mean speeds less than 30 km/hr and S_{85} less than 36 km/hr, whereas wider ones experienced mean speeds from 40 to 46 km/hr and S_{85} from 45 to 52 km/hr.

Because of the powerful effect of width on speed, it is difficult to evaluate, without holding width constant, the importance of street length in relation to speed. The correlation between length and 85th-percentile speeds was found to be .58, and the relationship of length to mean speed was found to be slightly less (.55). One of the wider (effective lane = 7.8 m) and shorter (length = 400 m) streets experienced speeds that appear lower than what might be expected, considering the rest of the pattern of speed versus width. It is possible that speeds are lower when width and length are less and are higher when both are larger. However, further attempts to explore the combined effects of street length and width on speed must be made before any conclusions can be inferred.

The correlation between speed and parking density was found to be .51 when the average speed was used and .56 for the 85th-percentile speed ($p < .01$). It is possible that the absence or presence of on-street parking is a more relevant characteristic than parking density and that better correlations might be found between speed and this latter measure. All streets selected for study had parking on both sides.

The correlations found between speed and alignment were also less significant than those between speed and traffic lane width. The correlation with the 85th-percentile speed was slightly higher than that with the average speed ($r = .40$, $p < .01$). It is possible that reduced sight distances due to curved alignment are more likely to produce a meaningful reduction in traffic speeds on two-way streets than on one-way streets, where the potential conflicts are less obvious. Only one-way streets were selected for this study. Another possible explanation is that alignment on the streets selected was not sufficiently severe to affect speeds significantly.

ENVIRONMENTAL QUALITY ASSESSMENT

In the final section of this paper, subjective ratings developed in this study are compared with subjective ratings developed in past work by other authors. In the search for valid methods to determine the environmental quality of streets, a variety of approaches have been used, and sets of criteria have been proposed, most of them objectively oriented. Apel (*15*) and Topp (*16*) are among the researchers who have studied this subject.

The model proposed by Apel estimates the multiple effects of traffic on the environment as a function of four main parameters, all of which are capable of being measured and quantified: noise, vehicle emissions (fumes), severance and pedestrian risk, and visual character of the street's cross section.

Topp suggested a methodology for determining street environmental capacity. It can be said that the procedure is based on a tradeoff between street appearance and traffic volume and six other main factors concerning street use by both pedestrians and traffic. A measure of the environmental quality of the street can be obtained by comparing the actual traffic volume in a street with its environmental capacity (optimum traffic volume).

In this study, the problem of assessing environmental quality is approached subjectively. In fact, it is the perceived environmental quality that is being assessed, defined in terms of user satisfaction and determined through attitudinal studies.

In order to assess environmental quality, the weighted sum of residents' average ratings of the street was used as an index. The overall resident satisfaction with the street environment was previously assumed to be a function of the level of satisfaction (negative or positive perception) with individual street attributes:

$$S_{jk} = \sum_{i=1}^{n} c_i \left(S_{ijk}\right)$$

where

S_{jk} = overall satisfaction level of resident j on street k;
$i = 1, 2, 3, \ldots, n$ street attributes;
S_{ijk} = satisfaction level with attribute i of resident j on street k; and
c_i = relative weight of attribute i in resident's overall evaluation.

On this basis, environmental quality may be assessed and expressed as the weighted sum of average user (resident) satisfaction levels (positive or negative perceptions) with street attributes, so that

$$E_k = \sum_{i=1}^{n} c_i \overline{S}_{ik}$$

where E_k is environmental quality of street k, and

$$\overline{S}_{ik} = \frac{1}{m} \sum_{j=1}^{m} \overline{S}_{ijk} \qquad j = 1, 2, 3, \ldots, m \text{ residents}$$

The streets in this study were ranked according to environmental quality so that the results obtained (subjective assessment) could be compared with those obtained by two other assessment models developed in Europe by Unterman (*14*) and by Apel (*15*). The Kendall coefficient of concordance was used to assess the correlation between street rankings, which yielded a fair agreement ($W = 0.73$).

CONCLUSIONS AND DISCUSSION

The main conclusions reached in this study are discussed as follows:

1. The residents' perceptive structure of street attributes was provided by four categories of correlated variables (extracted by subjecting the questionnaire data to a prin-

cipal components analysis), defined as follows:

- Traffic safety,
- Aesthetics and comfort,
- Sense of street community, and
- Traffic nuisance.

2. Individual perception of traffic safety can be satisfactorily predicted on the basis of environmental and traffic variables within the range measured. A number of explanatory variables were revealed; street curvature, steepness, and traffic speed were found to be negatively associated with individual perception. In situations where traffic flow is relatively light, residents' perception is most affected by traffic speed. The impact of speed may be modified by the street alignment: the steeper and more curved the street, the more traffic will be perceived as fast and dangerous.

3. Individual perception of street aesthetics and comfort can be satisfactorily predicted on the basis of environmental and traffic variables within the range focused on in the study. Residents' response to the components of the street view was found to depend greatly upon the width of the road visible, the number of on-street parked cars, and the degree of street enclosure. Comfort and visual satisfaction increase with higher degrees of enclosure, narrower traffic lanes, and lower parking density.

4. The results can have important implications for street design and traffic management. Reducing lane width is effective both in slowing down vehicles and in improving visual appearance. Similarly, creating visual constrictions by defining parking spaces and alternating them with strategically located landscaped nodes can both reduce traffic speeds and improve street space enclosure. These measures may be preferable to other measures for reducing traffic speeds, such as bending lane alignment and reducing sight distances, because it has been shown that these are not likely to produce any improvement in residents' sense of safety unless a proper balance between curvature increase and speed reduction is achieved.

5. Environmental quality was assessed in each study street in terms of user satisfaction (expressed as the weighted sum of average residents' ratings of street attributes), and also according to two objectively oriented procedures developed in the Federal Republic of Germany (Table 3). The correlation between street rankings obtained with the different procedures yielded a fair degree of agreement, indicating that the subjective procedure is capable of assessing environmental quality at least as effectively as the objective ones.

6. The current study was exploratory in nature. It covered a wide range of street attributes and studied their interrelationships. The importance of the findings is, as yet, hampered by the limited amount of data collected. In terms of street attributes, the study was small in scale—all 10 streets selected fell into a fairly narrow, homogeneous range. It has been suggested that the study be extended to include a wider range of streets and different neighborhoods, cities, and street functions—perhaps eventually to different countries. In this way, the robustness of the findings can be assessed and eventually extended to design recommendations for improved street quality.

TABLE 3 STREET RANKINGS ACCORDING TO ENVIRONMENTAL QUALITY

	Ranking According to		
Street	Subjective Assessment	Apel (*15*)	Topp (*16*)
Zidquiahu	1	2	2
Nachshon	2	1	1
Disraeli	3	3	4
Hatzalafim	4	9	3
Qiriat Sefer	5	4	6
Ruth	6	8	7
Oren	7	10	9
Margalit	8	5	5
Yannay	9	6	8
Einstein	10	7	10

REFERENCES

1. A. Rapoport. *Human Aspects of Urban Form: Towards a Man-Environment Approach to Urban Form and Design.* Pergamon Press, London, 1977.
2. J. Morton-Williams, B. Hedges, and E. Fernando. *Road Traffic and the Environment.* Social and Community Planning Research, London, 1978.
3. D. W. Robinson. *An Outline Guide to Criteria for the Limitations of Urban Noise.* Acro Report AC 39. National Physical Laboratory of Israel, Jerusalem, 1969.
4. L. L. Beranek. *Acoustics.* McGraw-Hill, New York, 1954.
5. A. Wilson. *Noise: Final Report.* Her Majesty's Stationery Office, London, 1962.
6. A. Alexandre, J. Barde, C. Lamure, and F. J. Langdon. *Road Traffic Noise.* Applied Science Publishers, London, 1975.
7. *Calculation of Road Traffic Noise.* U.K. Department of Environment, London, 1975.
8. D. C. Hothershall and R. J. Slater. *Transport and the Environment.* Crosby, Lockwood, Staples, London, 1977.
9. R. G. Hopkinson. The Evaluation of Visual Intrusion in Transport Situations. *Traffic Engineering and Control,* Vol. 14, No. 8, 1972, pp. 387–395.
10. R. E. Brindle. Residential Area Planning for Pedestrian Safety. Presented at Joint Australian Road Research Board/DOT Pedestrian Conference, Sydney, Australia, 1978.
11. B. Miedema, H. Menkerhorst, and H. H. van der Molen. A Model for the Subjective Experiencing of Traffic Safety in Residential Areas. Presented at Conference on Safety in the Built Environment, Portsmouth, U.K., July 1988, 13 pp.
12. K. Lynch. *Site Planning.* Massachusetts Institute of Technology Press, Cambridge, 1971.
13. J. Jacobs. *The Death and Life of Great American Cities.* Penguin Books, New York, 1961.
14. R. K. Unterman. *Accommodating the Pedestrian: Adjusting Towns and Neighborhoods for Walking and Bicycling.* Van Nostrand Reinhold, New York, 1984.
15. D. V. Apel. *Kraftverkehr und Umweltqualität von Stadtstrassen.* Kohlhammer Publishers, Stuttgart, West Germany, 1973.
16. H. H. Topp. Umfeldverträgliche Verkehrsbelastbarkeit städtischen Strassen: Ein kompensatorischen Ansatz. *Strasse und Autobahn,* Vol. 11, 1984.

Bayesian Identification of Hazardous Locations

JULIA L. HIGLE AND JAMES M. WITKOWSKI

A Bayesian analysis of accident data is used in the identification of hazardous locations. The Bayesian model used in the analysis is developed and discussed. Empirical comparisons of the results from the Bayesian analysis and from classical statistical analyses are also included. These comparisons suggest that there is an appreciable difference among the various identification techniques and that some classically based statistical techniques may be prone to err in the direction of false negatives.

One problem of ongoing interest in highway safety analysis is the identification of hazardous locations on the basis of historical data. Typically, a site is deemed hazardous if its recent accident history exceeds some specified level. One of the most common methods used in practice is to identify a site as hazardous if its accident rate over some period of time exceeds the mean accident rate over all sites in the region plus a multiple of the standard deviation of the site accident rates within that region over the same period of time. Such methods are based on the concept of confidence intervals within the context of classical statistics. The multiple used depends on the degree of confidence desired. Another commonly used technique is the rate-quality method (*1, 2*), which is based on statistical quality control procedures. This technique is used to calculate a critical accident rate, which depends on the degree of confidence desired, for each location. With the rate-quality method, a site is identified as hazardous if its observed accident rate exceeds its critical rate.

It is commonly acknowledged that because of the random variations that are inherent in accident phenomena, historical accident data do not always reflect long-term accident characteristics accurately. A site with a low accident rate (i.e., in the long run) may still have a high accident rate over a short period of time, and vice versa. Thus, the identification of hazardous locations is an inexact science at best. Regardless of the identification method used, traffic analysts will generally agree that the accident rate associated with a particular site is a random variable, a quantity that cannot be predicted with absolute certainty. Moreover, although regional accident characteristics may provide some useful information regarding the accident rate at a particular site, each site must be evaluated separately and should only be compared with sites that have similar underlying characteristics. The vast differences in accident histories that one finds among various sites suggest that the random variables used to describe the accident rates should differ from site to site.

To overcome some of the difficulties associated with the identification of hazardous locations, researchers have increasingly advocated the use of Bayesian analysis in this identification process (*3–7*). Bayesian analysis provides a framework wherein regional accident characteristics can be combined with site-specific accident histories, which results in a coherent method by which the random variables representing the accident rates at the various sites can be mathematically defined. Moreover, by using a Bayesian identification technique, one can identify hazardous sites on the basis of the probability that the accident rate exceeds some level. Such probabilistic identification methods differ both qualitatively and quantitatively from the confidence-based identification methods.

The research reported in this paper can be viewed as a complement to the research presented by Hauer and Persaud (*4–7*), although the techniques used differ substantially. These earlier papers are concerned with predicting the number of accidents that will occur at a particular location, and our research revolves around the accident rate at a particular location. An accurate prediction of the number of accidents at a particular site is invaluable in the assessment of the effectiveness of an improvement program, especially when one considers the phenomenon of regression to the mean. However, before an improvement program is implemented, one must first decide which sites require improvement. The contribution of this paper lies in the identification phase of the improvement process. Specifically, we develop a method for identifying hazardous locations on the basis of a Bayesian analysis of the accident data.

In this paper, we present the results of a Bayesian analysis of accident data from the jurisdiction of the Pima County Department of Transportation in Tucson, Arizona. The paper is divided into a discussion of the Bayesian methodology used in the study, a description of the data used, a comparison of the results of our Bayesian analysis with the results of an analysis based on classical statistical techniques, and our conclusions.

J.L. Higle, Department of Systems and Industrial Engineering, University of Arizona, Tucson, Ariz. 85721. J.M. Witkowski, JHK and Associates, 111 S. Church Street, Suite 470, Tucson, Ariz. 85701.

BAYESIAN METHODOLOGY

Bayesian analysis differs significantly from the classical statistical analysis of accident data. The motivation for the use of the Bayesian analysis is the desire to treat the actual accident rate (i.e., the number of accidents per million vehicles entering an intersection) at a particular location as a random variable and to use a combination of the regional accident characteristics and the accident history at that location to determine the probability that the location is hazardous. In this way, we hope to better utilize the available information throughout the identification process.

Our Bayesian analysis uses a two-step procedure. In the first step, we aggregate the accident histories across a number of sites (i.e., across all sites within an appropriately defined region). The result of this step is a gross estimation of the probability distribution of the accident rates across the region. We then use this regional distribution and the accident history at a particular site to obtain a refined estimation of the probability distribution associated with the accident rate at that particular site. Naturally, we obtain this refined estimation for all sites within the region, and two sites with equivalent histories will have identically refined distributions. This essentially concludes the Bayesian portion of the analysis. With the collection of refined distributions, one can now assess the probability that any given site is hazardous.

To formally describe the Bayesian identification process, we require the following notation:

$\tilde{\lambda}_i$ = accident rate at location i (note that $\tilde{\lambda}_i$ is treated as a random variable);
N_i = number of accidents at location i during the period of time in question;
V_i = number of vehicles passing through location i during the period of time in question;
$f_i(\lambda \mid N_i, V_i)$ = probability density function associated with the accident rate at location i, given the observations N_i and V_i; and
$f_R(\lambda)$ = probability density function associated with the accident rate across the region.

Thus, $f_R(\lambda)$ represents the gross estimation of the probability distribution of the accident rate across the region, and $f_i(\lambda \mid N_i, V_i)$ represents the refined estimation of the probability distribution at site i, as previously discussed. Moreover, the cumulative distribution function associated with the accident rate, $\tilde{\lambda}_i$, is given by

$$P\{\tilde{\lambda}_i \le \hat{\lambda}\} = \int_0^{\hat{\lambda}} f_i(\lambda \mid N_i, V_i)\, d\lambda$$

In performing this analysis, we make the following assumptions, which are similar to those of Morin, Norden et al., Hauer and Persaud, and Glauz et al. (*1, 2, 4, 6–8*), to name but a few.

A1. At any given location, when the accident rate is known (i.e., if $\tilde{\lambda}_i = \lambda$), the actual number of accidents follows a Poisson distribution with expected value λV_i. That is,

$$P\{N_i = n \mid \tilde{\lambda}_i = \lambda, V_i\} = \frac{(\lambda V_i)^n}{n!} e^{-\lambda V_i}$$

A2. The probability distribution of the regional accident rate, $f_R(\lambda)$, is the gamma distribution.

The first assumption indicates that because the actual accident rate is explicitly treated as a random variable, the conditional distribution of the number of accidents (given the accident rate) is the Poisson distribution. The second assumption implies that

$$f_R(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

for some α and β. Thus, the first step associated with the Bayesian analysis, that of determining $f_R(\lambda)$, is equivalent to determining the values of α and β. There are a number of possibilities.

The most commonly used estimates are the method of moments estimates (MME), where α and β are chosen so that the mean and variance associated with the gamma distribution are equal to the mean and variance of the sample. That is, let $\bar{x}$ be the sample mean of the observed accident rates, s^2 be the sample variance of the observed accident rates, and m be the number of sites in the region. Then

$$\bar{x} = \frac{1}{m} \sum_{i=1}^{m} \frac{N_i}{V_i}$$

$$s^2 = \frac{1}{m-1} \sum_{i=1}^{m} \left(\frac{N_i}{V_i} - \bar{x}\right)^2$$

Using the MME, one selects α and β so that $\bar{x} = \alpha/\beta$ and $s^2 = \alpha/\beta^2$, or equivalently, $\beta = \bar{x}/s^2$ and $\alpha = \beta\bar{x}$.

Other commonly used estimates are the maximum likelihood estimates (MLE), where α and β are chosen so that they represent the values that are most likely to have generated the observed data. That is, if $\hat{\lambda}_i$ is the observed accident rate at site i (i.e., $\hat{\lambda}_i = N_i/V_i$), then α and β are chosen to maximize

$$\mathscr{L}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m \mid \alpha, \beta) = \prod_{i=1}^{m} \frac{\beta^\alpha}{\Gamma(\alpha)} \hat{\lambda}_i^{\alpha-1} e^{-\beta\hat{\lambda}_i}$$

$$= \left\{\frac{\beta^\alpha}{\Gamma(\alpha)}\right\}^m \left[\prod_{i=1}^{m} \hat{\lambda}_i\right]^{\alpha-1} e^{-\beta \sum_{i=1}^{m} \hat{\lambda}_i}$$

The function $\mathscr{L}$ represents the likelihood function associated with the observed data when the parameters α and β are assumed. The MLE values for α and β may be obtained by solving the equations

$$\frac{\partial \mathscr{L}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m \mid \alpha, \beta)}{\partial \alpha} = 0$$

$$\frac{\partial \mathscr{L}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m \mid \alpha, \beta)}{\partial \beta} = 0$$

Although the MME and the MLE are among the most commonly used methods of parameter estimation, other methods for estimating α and β exist and are discussed at length by Berger (*9*).

Once values for α and β have been determined, the first step of the analysis has been completed. In the second step, the observed accident rate at each site is used in combination with the gross estimate of the regional probability distribution to obtain the site-specific probability density functions, $f_i\,(\lambda \mid N_i, V_i)$. These density functions are obtained using Bayes's theorem. That is,

$$f_i(\lambda \mid N_i, V_i) \propto f(N_i \mid \lambda, V_i) f_R(\lambda)$$

Within the framework of Bayesian analysis, it is well known that under Assumptions A1 and A2, the resulting probability distribution $f_i(\lambda \mid N_i, V_i)$ is a gamma distribution (*9, 10*). Moreover, the parameters associated with this distribution, α_i and β_i, are easily obtained from the original choices of α and β and the observed data, N_i and V_i, as follows:

$$\alpha_i = \alpha + N_i$$

$$\beta_i = \beta + V_i$$

Thus, the probability density function associated with the accident rate at location $i(\tilde{\lambda}_i)$ is given by

$$f_i(\lambda \mid N_i, V_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i \lambda}$$

Note that as N_i and V_i increase, the site-specific parameters (α_i and β_i) will be largely determined by the observed data (N_i and V_i) and will become insensitive to the initial choice of α and β. As such, for each computation, it may be preferable to use the MME values rather than the MLE values, because they are substantially easier to calculate. All computations within this paper were based on the MME values of α and β.

With this collection of probability density functions, the identification of hazardous locations is now a straightforward matter. If $\overline{\lambda}$ is an upper limit on the "acceptable" accident rates, then we wish to identify a site i as hazardous if the probability is significant that $\tilde{\lambda}_i$ exceeds $\overline{\lambda}$. That is, if

$$P(\tilde{\lambda}_i > \overline{\lambda} \mid N_i, V_i) > \delta$$

where δ is some predetermined tolerance level, then site i is recognized as a hazardous location. Naturally, the appropriate values for $\overline{\lambda}$ and δ must be determined. For example, in the results section of this paper, various values of $\overline{\lambda}$ and δ are used to develop criteria for the identification of hazardous locations that are analogous to the criteria used in classically based statistical procedures. This allows a direct comparison between the results obtained from the Bayesian procedure presented in this paper and the results obtained from the classical techniques.

DATA DESCRIPTION

For the purposes of this study, 5-year (July 1981–June 1986) accident histories for signalized intersections under the jurisdiction of the Pima County Department of Transportation, in Tucson, Arizona, were used. Because significant improvement plans were undertaken during the third year (July 1983–June 1984), the data were broken into two separate sets. The first set corresponds to July 1981–June 1983, whereas the second set corresponds to July 1984–June 1986. Between July 1981 and June 1984, four intersections were signalized. Thus, the first data set includes 33 intersections, and the second data set includes 37 intersections. The two data sets were analyzed independently. For each intersection, the observed accident rate was calculated as the ratio of the total number of accidents to the total traffic volume over the 2-year period. The data are summarized in Tables 1 and 2.

One should note that the observed accident rate over each 2-year period is calculated as $N \times 10^6/2V \times 365$, and thus is normalized to represent the accident rate per million vehicles entering the intersection. The last two columns represent the probability that the site is hazardous on the basis of the two criteria developed in the section on results; these elements are discussed in further detail in that section.

EMPIRICAL RESULTS

In order to compare the results of an analysis based on classical statistical methods (e.g., those based on statistical confidence intervals) with the results of an analysis based on the Bayesian methodology, the two analyses must use analogous criteria in identifying a hazardous location. In practice, two commonly used criteria can be stated as follows.

C1. Site i is hazardous if the observed rate, $\hat{\lambda}_i$, exceeds the observed average rate across the region, x, with a level of confidence equal to δ.

C2. Site i is hazardous if the observed accident rate, $\hat{\lambda}_i$, exceeds the site's critical rate, which is a function of the observed regional accident rate, the traffic volume at site i, and the level of confidence desired, δ.

Typically, δ is a reasonably high number, such as 0.99, 0.95, or 0.90. C1 is the standard confidence-based crite-

TABLE 1 DATA OBTAINED FROM JULY 1981 THROUGH JUNE 1983

site number	observed accident rate (#/MVE)	number of accidents (N)	daily volume (V)	prob. (B1)	prob. (B2)
1	0.957	20	28644	0.4308	0.3861
2	1.192	46	52891	0.8684	0.8331
3	0.947	20	28950	0.4145	0.3701
4	1.437	43	40994	0.9813	0.9738
5	0.588	9	20965	0.0742	0.0614
6	1.043	14	18393	0.5402	0.5005
7	1.418	17	16422	0.8609	0.8377
8	0.779	9	15825	0.2573	0.2285
9	1.375	14	13953	0.8046	0.7776
10	1.007	18	24496	0.5058	0.4621
11	1.074	14	17863	0.5740	0.5349
12	1.174	21	24515	0.7280	0.6897
13	0.660	14	29054	0.0754	0.0608
14	1.040	22	28998	0.5632	0.5167
15	1.133	18	21773	0.6634	0.6237
16	0.675	11	22343	0.1198	0.1007
17	0.846	5	8100	0.3700	0.3411
18	0.742	11	20323	0.1897	0.1640
19	0.617	9	19989	0.0965	0.0810
20	0.282	4	19450	0.0061	0.0047
21	0.709	11	21253	0.1545	0.1318
22	1.003	8	10924	0.4805	0.4468
23	1.010	15	20360	0.5034	0.4626
24	0.088	1	15550	0.0025	0.0019
25	1.848	17	12605	0.9627	0.9543
26	0.567	6	14500	0.1138	0.0978
27	1.337	15	15376	0.7967	0.7683
28	1.471	46	42850	0.9891	0.9842
29	1.604	14	11957	0.8908	0.8727
30	1.032	15	19915	0.5311	0.4904
31	0.963	13	18502	0.4441	0.4054
32	1.184	20	23147	0.7308	0.6935
33	0.589	9	20953	0.0745	0.0616

rion, whereas C2 corresponds to the rate-quality criterion, developed by Norden et al. (*2*).

To identify hazardous locations by using Criterion C1, one must calculate both the sample mean $\bar{x}$ and the sample standard deviation s. Associated with each value of δ is a constant k_δ (e.g., $k_{0.95} = 1.645$), and if

$$\hat{\lambda}_i > \bar{x} + k_\delta s$$

then site i is said to be hazardous at the δ confidence level (*11*). To identify hazardous locations using the Bayesian methodology, a criterion that is analogous to C1 can be stated as follows:

B1. Site i is hazardous if the probability is greater than δ that its true accident rate, $\hat{\lambda}_i$, exceeds the observed average rate across the region.

Recall that the Bayesian methodology treats the accident rate at a particular location as a random variable and obtains a refined estimate of its probability distribution. As such, if

$$P\{\tilde{\lambda}_i > \bar{x} \mid N_i, V_i\} > \delta$$

then site i is said to be hazardous. Thus, the identification of hazardous locations using Criterion B1 involves the computation of

$$P\{\tilde{\lambda}_i > \bar{x} \mid N_i, V_i\} = 1 - P\{\tilde{\lambda}_i \leq \bar{x}\}$$
$$= 1 - \int_0^{\bar{x}} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i \lambda} d\lambda \qquad (1)$$

If the computed value exceeds δ, site i is identified as hazardous.

Similarly, to identify hazardous sites using Criterion C2, one must calculate the regional accident rate,

$$x_R = \frac{\sum_i N_i}{\sum_i V_i}$$

For a given level of confidence, δ, the critical rate associated with location i is computed as follows:

$$\lambda c_i = x_R + k_\delta \sqrt{\frac{x_R}{V_i}} + \frac{1}{2V_i}$$

TABLE 2 DATA OBTAINED FROM JULY 1984 THROUGH JUNE 1986

site number	observed daily rate (#/MVE)	number of accidents (N)	daily volume (V)	prob. (B1)	prob. (B2)
1	0.710	15	28950	0.0727	0.0616
2	1.006	51	69450	0.3890	0.3405
3	1.047	27	35350	0.4845	0.4458
4	0.897	22	33600	0.2398	0.2118
5	1.051	23	30000	0.4885	0.4522
6	0.625	12	26300	0.0414	0.0347
7	0.875	16	25050	0.2443	0.2190
8	0.726	11	20750	0.1266	0.1114
9	0.865	13	20600	0.2544	0.2303
10	0.648	13	27500	0.0463	0.0388
11	1.028	18	24000	0.4520	0.4191
12	1.357	36	36350	0.9083	0.8903
13	1.237	31	34350	0.7828	0.7528
14	1.162	24	28300	0.6560	0.6222
15	0.704	15	29200	0.0681	0.0575
16	0.689	14	27850	0.0648	0.0548
17	0.857	6	9592	0.3238	0.3033
18	1.109	27	33350	0.5909	0.5533
19	1.318	28	29100	0.8436	0.8202
20	0.386	8	28400	0.0017	0.0013
21	1.342	29	29600	0.8669	0.8456
22	1.269	15	16200	0.7032	0.6770
23	0.947	18	26050	0.3340	0.3036
24	0.684	12	24050	0.0790	0.0679
25	2.289	34	20350	0.9998	0.9997
26	1.006	13	17700	0.4275	0.3984
27	1.300	24	25300	0.8034	0.7782
28	2.177	66	41550	1.0000	1.0000
29	1.912	30	21500	0.9951	0.9937
30	1.288	22	23400	0.7795	0.7534
31	0.988	16	22200	0.3983	0.3675
32	1.266	20	21650	0.7445	0.7173
33	1.160	26	30700	0.6630	0.6284
34	0.601	5	11400	0.1361	0.1230
35	0.813	11	18550	0.2145	0.1935
36	0.525	12	31300	0.0081	0.0064
37	0.602	9	20500	0.0569	0.0489

This critical rate is based on the assumption that the number of accidents at location i is Poisson distributed with a mean of $x_R V_i$ (*2*), which is similar to Assumption A1 in this paper. The critical rate is defined so that with this assumption, the observed accident rate will be less than or equal to the critical rate with probability δ. An investigation into the early development of this rate-quality method (*2*) suggests that an analogous criterion within the Bayesian methodology can be stated as follows:

B2. Site i is hazardous if the probability is greater than λ that its accident rate, $\tilde{\lambda}_i$, exceeds the observed regional accident rate, x_R.

That is, under Criterion B2, site i is identified as hazardous if

$$P\{\tilde{\lambda}_i > x_R \mid N_i, V_i\} > \delta$$

or equivalently, if

$$1 - \int_0^{x_R} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i - 1} e^{-\beta_i \lambda} d\lambda > \delta \qquad (2)$$

Note that the assumptions leading to Criteria C2 and B2, namely, those regarding the Poisson nature of accidents at a particular site, are very similar. The fundamental difference lies in the fact that in using C2, one implicitly assumes that the true accident rate is x_R (*2*). The authors who pioneered this method concede that the true rate (*2*) "is never known and we shall always have to be satisfied with an estimate of the expectation" (i.e., x_R). In using Criterion B2, one accounts for the inherent randomness associated with each accident rate, as reflected in Assumption A2.

In identifying hazardous locations on the basis of the Bayesian methodology (i.e., Criteria B1 and B2), one must perform the integrations identified in Equations 1 and 2. A computer program was written to numerically evaluate each of these integrals. The results of our empirical study are summarized in Figures 1 through 4. For each data set and for each value of δ (i.e., $\delta = 0.99$, $\delta = 0.95$, and $\delta = 0.90$), hazardous sites were identified on the basis of Criteria C1 and C2, corresponding to the classical statistical methods, and of the analogous Bayesian Criteria B1 and B2. These results are presented in Figures 1–4. The ele-

a. $\delta = 0.99$

	HC1	NC1
HB1	0	0
NB1	0	33

b. $\delta = 0.95$

	HC1	NC1
HB1	1 (25)	2 (4,28)
NB1	1 (29)	29

c. $\delta = 0.90$

	HC1	NC1
HB1	2 (25,28)	1 (4)
NB1	1 (29)	29

FIGURE 1 Distribution of sites based on B1 and C1 (July 1981–June 1983).

a. $\delta = 0.99$

	HC2	NC2
HB2	0	0
NB2	2 (25,28)	31

b. $\delta = 0.95$

	HC2	NC2
HB2	3 (4,25,28)	0
NB2	0	30

c. $\delta = 0.90$

	HC2	NC2
HB2	3 (4,25,28)	0
NB2	2 (7,29)	28

FIGURE 2 Distribution of sites based on B2 and C2 (July 1981–June 1983).

a. $\delta = 0.99$

	HC1	NC1
HB1	2 (25,28)	1 (29)
NB1	0	34

b. $\delta = 0.95$

	HC1	NC1
HB1	3 (25,28,29)	0
NB1	0	34

c. $\delta = 0.90$

	HC1	NC1
HB1	3 (25,28,29)	1 (12)
NB1	0	33

FIGURE 3 Distribution of sites based on B1 and C1 (July 1984–June 1986).

a. $\delta = 0.99$

	HC2	NC2
HB2	3 (25,28,29)	0
NB2	0	34

b. $\delta = 0.95$

	HC2	NC2
HB2	3 (25,28,29)	0
NB2	0	34

c. $\delta = 0.90$

	HC2	NC2
HB2	3 (25,28,29)	0
NB2	1 (12)	33

FIGURE 4 Distribution of sites based on B2 and C2 (July 1984–June 1986).

ments in the 4×4 matrices found in Figure 1 are organized as follows:

1. Columns

a. HC1 corresponds to the number of sites that were identified as hazardous on the basis of Criterion C1 (i.e., the number of sites whose observed accident rate exceeded $\bar{x} + k_\delta s$).

b. NC1 corresponds to the number of sites that were not identified as hazardous on the basis of Criterion C1.

2. Rows

a. HB1 corresponds to the number of sites that were identified as hazardous on the basis of Criterion B1 (i.e., the number of sites with $P\{\tilde{\lambda}_i > \bar{x}\} \geq \delta$).

b. NB1 corresponds to the number of sites that were not identified as hazardous on the basis of Criterion B1.

The numbers in parentheses correspond to the sites that are identified as hazardous, as represented in Tables 1 and 2.

Thus, for example, in Figure 1 we see that for July 1981–June 1983, for $\delta = 0.95$, one site (25) is identified as hazardous under both C1 and B1, two sites (4, 28) are identified as hazardous under B1 but not C1, and one site (29) is identified as hazardous under C1 but not B1. The remaining 29 sites are not identified as hazardous under either criterion.

The results of the comparison between C2 and B2 for July 1981–June 1983 are similarly arranged in Figure 2 and those for the analyses of the data collected between July 1984 and June 1983 are presented in Figures 3 and 4 for B1 and C1 and for B2 and C2, respectively.

Because there seems to be consistent disagreement between the various criteria, a discussion of the information conveyed in Figures 1–4 is in order. First, note that as expected, as δ decreases, the number of sites identified increases under all four criteria. That is, the more relaxed the identification requirement, the easier it is to be identified.

Second, there is very little difference between the sites identified by B1 and B2, the Bayesian criteria. The only difference is in Site 12, using the data collected from July 1984–June 1986. With $\delta = 0.90$, it is identified using B1 but not B2 (see Figures 3 and 4). However, the data presented in Table 2 indicate that the probability computed under Criterion B1 is 0.9083, whereas the probability computed under Criterion B2 is 0.8901. Thus, although the methods differ, the difference is not substantial. This significant agreement between the two methods is easily explained by the data. B1 uses the threshold value $\overline{\lambda} = \bar{x}$, whereas B2 uses $\overline{\lambda} = x_R$. For both data sets, x and x_R are not substantially different, as indicated by the summary statistics presented in Table 3.

One should note carefully that Criteria B1, B2, and C2 all tend to be more conservative than C1, in that they tend to identify more sites as hazardous. This suggests that C1 may be more susceptible to the identification of false negatives (i.e., those sites that are actually hazardous but are not identified as such).

In a review of Figures 3 and 4, it is clear that for the data collected between July 1984 and June 1986, the classical criteria are in relatively high agreement with their Bayesian counterparts. This is most likely due to the extremely high accident rates (2.289, 2.177, and 1.912, respectively) of three intersections (25, 28, and 29), compared with a mean accident rate of 1.0396 and a regional rate of 1.0578. Because of the extreme nature of these three accident rates, it is reasonable to expect that any justifiable procedure would identify these sites as hazardous and that all others may seem safe by comparison. Of course, when the Bayesian procedure is used, a change in the threshold value would affect the sites that are identified as hazardous. For the purposes of this study, the values of $\overline{\lambda} = \bar{x}(\text{B1})$ and $\overline{\lambda} = x_R(\text{B2})$ were chosen so that the Bayesian and classical methods could be easily compared.

TABLE 3 SUMMARY STATISTICS

Data Set	$\bar{x}$	x_R	s
June 1981 - July 1983	0.9815	1.0042	0.3756
June 1984 - July 1986	1.0396	1.0578	0.4196

The disagreement between Criteria B1 and C1 (Figures 1 and 3) is probably best explained in terms of the underlying assumptions. B1 is based on the widely accepted assumptions that accidents occur according to a Poisson distribution and that the accident rate has a gamma distribution (i.e., Assumptions A1 and A2 as stated in the section on Bayesian methodology in this paper). Similarly, Criterion C1 is based on the implicit assumption that the observed accident rates are normally distributed. The relatively large standard deviations when compared with the low sample means (e.g., $\bar{x} = 0.9815$ and $s = 0.3756$ for the July 1981–June 1983 data) combined with the fact that the accident rates must be nonnegative suggest that the normal distribution may yield an inappropriate model for these data sets.

Because Criteria B2 and C2 are based on a similar assumption (i.e., that accidents occur according to a Poisson distribution), the differences between the corresponding results, as summarized in Figures 2 and 4, are due solely to the treatments of the actual accident rate. The Bayesian method explicitly assumes that the accident rate at any given site is a random variable and accounts for this randomness in the identification process. The rate-quality method (C2) implicitly assumes that the accident rate at each site is equal to the regional rate x_R. Thus, the Bayesian method allows site-specific accident information to guide the identification process, and the rate-quality method does not.

Finally, it should be noted that Intersection 4, which was identified as hazardous by using the data collected from July 1981–June 1983 with $\delta = 0.95$ under Criteria B1, B2, and C2, underwent significant change during July 1983–June 1984. Subsequently, it was no longer identified as a hazardous intersection, and the probability that it is hazardous dropped from 0.9813 to 0.2398 on the basis of B1 or from 0.9738 to 0.2190 on the basis of B2. Clearly, these substantial drops indicate that the improvement program was successful.

CONCLUSIONS

Use of a Bayesian analysis in the identification of hazardous accident locations using accident rate data appears to be a fundamentally sound procedure, which is shown to have identification criteria analogous to those used in the classical identification scheme, although it is certainly not limited to these criteria. The Bayesian technique has the added advantage of allowing the assessment of the impact of varying the degree of confidence, δ, without requiring that the decision statistics be recomputed. Moreover, knowing the probability that the actual rate exceeds the regional rate, for example, provides added information that can be used to evaluate the trade-offs involved in deciding which sites are candidates for improvement funds.

The results presented in Figures 1 and 3 suggest that, in general, the confidence-based procedure, C1, may be inappropriate for identifying hazardous locations. Criterion C1 fails to identify as hazardous many sites that are flagged by Criteria B1, B2, and C2. The underlying assumption of normality in the distribution of the accident rate appears to cause C1 to err in the direction of false negatives. This is the least desirable characteristic for an identification procedure.

The results presented in Figures 2 and 4, combined with an analysis of the underlying assumptions, suggest that in many cases, the use of B2 may be preferable to the use of C2. This may be especially true when data are sparse or when numerous years of comparable data are not available. It is expected that the differences between B2 and C2 will be substantially reduced whenever a great deal of data are available.

REFERENCES

1. D. A. Morin. Application of Statistical Concepts to Accident Data. In *Highway Research Record 187,* HRB, National Research Council, Washington, D.C., 1967, pp. 72–79.
2. M. Norden, J. Orlansky, and H. Jacobs. Application of Statistical Quality-Control Techniques to Analysis of Highway-Accident Data. In *Bulletin 117,* HRB, National Research Council, Washington, D.C., 1956, pp. 17–31.
3. C. P. Brinkman. Toward Improved Highway Safety Analysis. *Public Roads,* Vol. 50, No. 2, 1986, pp. 49–54.
4. E. Hauer. On the Estimation of the Expected Number of Accidents. *Accident Analysis and Prevention,* Vol. 18, No. 1, 1986, pp. 1–12.
5. E. Hauer and B. Persaud. Common Bias in Before-and-After Accident Comparisons and Its Elimination. In *Transportation Research Record 905,* TRB, National Research Council, Washington, D.C., 1983, pp. 164–174.
6. E. Hauer and B. Persaud. The Problem of Identifying Hazardous Locations Using Accident Data. In *Transportation Research Record 905,* TRB, National Research Council, Washington, D.C., 1984, pp. 36–43.
7. B. Persaud and E. Hauer. A Comparison of Two Methods for De-Biasing Before-and-After Accidents Studies. In *Transportation Research Record 975,* TRB, National Research Council, Washington, D.C., 1984, pp. 43–49.
8. W. D. Glauz, K. M. Bauer, and D. J. Migletz. Expected Traffic Conflict Rates and Their Use in Predicting Accidents. In *Transportation Research Record 1026,* TRB, National Research Council, Washington, D.C., 1985, pp. 1–12.
9. J. O. Berger. *Statistical Decision Theory and Bayesian Analysis,* 2nd ed. Springer Verlag, New York, 1985.
10. M. D. DeGroot. *Optimal Statistical Decisions.* McGraw-Hill, New York, 1970.
11. J. C. Laughlin, L. E. Haefner, J. W. Hall, and D. R. Clough. *NCHRP Report 162: Methods for Evaluating Highway Safety Improvements.* TRB, National Research Council, Washington, D.C., 1975.

DISCUSSION

CARL N. MORRIS
Department of Mathematics and Statistical Sciences Center, University of Texas, Austin, Tex. 78712

The proposal by Higle and Witkowski to use hierarchical Bayesian and empirical Bayesian methods to identify hazardous highway locations is a sound one that can be expected to fare well in comparison with more traditional methods. Accident rate estimation is extremely uncertain because the number of accidents at any one intersection tends to be quite random and subject to the regression-to-the-mean phenomenon. Bayesian methods help because they effectively permit pooling of data from other relevant sites. It also makes good sense that the authors base hazard determinations on probabilistic assessments of the value $\tilde{\lambda}_i$ of the intrinsic hazard rate, as, for example, Criteria B1 and B2 require.

In Table 1 the exposure rates (daily volumes) V_i vary by a factor of 6.5 between the two extreme sites (Sites 2 and 17). Although the setting here is for Poisson data, this is otherwise analogous to the empirical Bayes estimation of means in the normal distribution case, which is now well understood (*1*, *2*). In such cases the Bayesian ranking of extreme intersections differs from the observed daily rates because those extreme intersections with low volumes generally would regress to the mean more than high-volume intersections. Thus, Site 4 in Table 1 probably is more hazardous than Site 25, even though the observed rate of 1.437 MVE is less than 1.848 MVE for Site 25. This occurs because the daily volume for Site 4 is more than three times higher, 40,994 to 12,605.

It should be obvious that a methodology that properly weighs all evidence in ranking dangerous intersections is very valuable. Such features can only be revealed by using Bayesian and empirical Bayesian models, in which distributions are specified and estimated for both observed data and unobserved parameters.

Despite the virtues of the authors' general idea, there are features in their proposed methods that need further adjustment to correct for bias and to improve statistical efficiency. I will explain the difficulties partly on the basis of my own research with Olga Pendleton on accident analysis using these same Poisson-gamma models.

The authors' development fits within the General Model for Statistics (*2*, *3*), which in their Poisson-gamma setting and notation is summarized in Table 4.

Note that Equations 3 and 4 correspond to the authors' Assumptions A1 and A2. The descriptive model is entirely equivalent to the inferential model (Table 5), which reverses the probabilistic conditioning and is more convenient for statistical analysis.

Expressions 5 and 7 refer to the negative binomial and gamma distributions, with their usual parameterization, whereas Expressions 6 and 8 and the square brackets signify that the means and variances are displayed.

TABLE 4 DESCRIPTIVE MODEL

Observed Data:	$N_i \mid \tilde{\lambda}_i \sim \text{Poisson}(V_i\tilde{\lambda}_i)$ $i = 1, \ldots, m$ independently	(3)
Unobserved Parameters:	$\tilde{\lambda}_i \mid (\alpha,\beta) \sim \text{Gamma}\left(\alpha, \frac{1}{\beta}\right)$ $i = 1, \ldots, m$ independently $\phi \equiv (\alpha,\beta)$ unknown, $\alpha,\beta > 0$	(4)

TABLE 5 INFERENTIAL MODEL

Observed Data:	$N_i \mid \alpha, \beta \sim \text{NegBin}\left(\alpha, p_i \equiv \frac{V_i}{V_i + \beta}\right)$	(5)
	$= \text{NegBin}\left[\frac{\alpha V_i}{\beta}, \frac{\alpha V_i}{\beta} + \frac{\alpha V_i^2}{\beta^2}\right]$	(6)
	$i = 1, \ldots, m$ independently.	
Unobserved Parameters:	$\tilde{\lambda}_i \mid N_i, \alpha, \beta \sim \text{Gamma}\left(\alpha + N_i, \frac{1}{\beta + V_i}\right)$	(7)
	$= \text{Gamma}\left[\frac{\alpha + N_i}{\beta + V_i}, \frac{\alpha + N_i}{(\beta + V_i)^2}\right]$	(8)
	$i = 1, \ldots, m$ independently.	

The analysis proceeds using (N_i, . . . , N_m) to estimate $\phi = (\alpha, \beta)$ from Expression 5, and then carries this information to 7 or 8 to assess the posterior distribution. The simplest method for doing this, often called "empirical Bayes," simply develops a point estimate $\hat{\phi} = (\hat{\alpha}, \hat{\beta})$ and substitutes these values into 7 or 8. [This can be risky if ($\hat{\alpha}$, $\hat{\beta}$) are not accurately estimated, an issue that could be assessed in the data example.] The authors follow this empirical Bayes approach, although not quite correctly.

Note that the marginal distribution (Expression 5) for the data N_i is negative binomial, not gamma, as the authors indicate when discussing the MLE. Thus, the maximum likelihood conditions in the second section of their paper are incorrect.

Similarly, the MME technique is improperly applied by the authors, and the estimates of the "hyperparameter" $\phi = (\alpha, \beta)$ are biased and inefficient. Define $X_i = N_i/V_i$, $X = 1/m\Sigma X_i$ and $s^2 = [1/(m - 1)]\Sigma(X_i - \bar{X})^2$, as in the second section. Then from Expression 6, $EX_i = \alpha/\beta$ and so $E\bar{X} = \alpha/\beta$, as claimed. However, from Expression 6,

$$\text{Var}(X_i) = \frac{\alpha}{\beta V_i} + \frac{\alpha}{\beta^2} \equiv \sigma^2 \tag{9}$$

and so

$$Es^2 = \frac{1}{m} \sum_1^m \sigma_i^2 \tag{10}$$

$$= \frac{\alpha}{\beta^2} + \frac{\alpha}{\beta V^*} \tag{11}$$

This exceeds α/β^2, the value claimed by the authors, by the amount $\alpha/\beta V^*$, V^* the harmonic mean of (V_i, . . ., V_m). It follows that, instead of the authors' formula, the MME is

$$\hat{\beta} = \frac{V^*\bar{X}}{V^*s^2 - \bar{X}} \qquad \hat{\alpha} = \bar{X}\hat{\beta} \tag{12}$$

Of course additional modifications are required if the denominator of β is not positive or is close to zero.

For the data set of Table 1 (1981–1983), the authors' formulas give values for α and β) of about 7 and 7, respectively, but the correct MME equations give estimates closer to 14 and 14. Thus, the model in Expression 4 should provide about twice as much information as the authors have estimated.

We could further improve the estimation. The unweighted mean $\bar{X}$ is not the best estimate of α/β because those X_i based on larger exposures V_i deserve more weight. A similar statement would apply to the use of s^2, but determining the correct weights is quite complicated in that case. The benefits of these improvements with m = 33 or 37, as in the examples, should be of second order, however, and so perhaps this use of a simple method is not costly for the data considered.

A nice feature of using the optimally weighted $\bar{X}$ is that it then could be used instead of $\bar{X}$ and X_R in Criteria B1 and B2 and would be preferable to either.

The simple device of substituting the estimates (α, β) from Expression 5 into 7 fails to acknowledge the uncertainty in knowledge of (α, β). More accurate methods, which are hard to derive [see discussion by Morris (*3*)] would spread out the posterior distribution. The effect of properly accounting for this in the Higle-Witkowski application would be to lower somewhat the probabilities for Criteria B1 and B2 for those locations with high accident rates.

To summarize, the Higle-Witkowski Bayesian model promises to have many advantages over standard methods for identifying hazardous locations. It will take more time before the most appropriate analytical methods are available, however. The development of such methods promises to be a rewarding and interesting task.

ACKNOWLEDGMENT

This discussion was based on research supported by FHWA at the Texas Transportation Institute, Texas A&M University, College Station, Texas. Olga Pendleton is the principal investigator.

REFERENCES

1. B. Efron and C. N. Morris. Data Analysis Using Stein's Estimator and Its Generalizations. *Journal of the American Statistical Association,* Vol. 70, 1975, pp. 311–319.
2. C. N. Morris. Parametric Empirical Bayes Inference: Theory and Applications (with discussion). *Journal of the American Statistical Association,* Vol. 78, 1983, pp. 47–65.
3. C. N. Morris. Determining the Accuracy of Bayesian Empirical Bayes Estimators in the Familiar Exponential Families. In *Statistical Decision Theory and Related Topics, IV,* Vol. 1 (S. S. Gupta and J. O. Berger, eds.), Springer-Verlag, New York, 1988, pp. 251–263.

The opinions expressed in this discussion are those of the author alone, and have not been reviewed by FHWA.

DISCUSSION

OLGA J. PENDLETON
Texas Transportation Institute, College Station, Tex. 77843.

Morris, in his discussion of this paper, has eloquently and didactically covered its weaknesses. The following discussion is offered to paraphrase and reemphasize some of his comments and to add a few comments regarding the numerical example.

Basically, this study uses a method that is inefficient. Furthermore, there are computational errors in the analysis. The authors' estimates are biased and inefficient and the method of moments technique is improperly applied, as Morris has noted. With regard to the computational errors, Morris has shown that the discrepancy in the parameter estimates is twofold when the computations are done correctly.

My second comment pertains to the numerical example. Although the results are computationally incorrect, I will base my comments on the data as they were originally presented, to show that many of the authors' claims are not supported by their numerical example.

The authors attempt to show by the numerical example that the empirical Bayesian (EB) methods (Criteria B1 and B2) are superior to two classical methods—one that assumes a normal distribution (C1) and one that assumes the more correct Poisson distribution (C2). Careful inspection of the results (Figures 1–4) does not support this claim.

In Figure 1 the classical estimate, C1, identifies one site as hazardous that the EB method fails to recognize (Site 29), and fails to identify one site that the EB method does identify (Site 4). Note that the accident rate for the site identified by the classical method is higher (1.604) than that identified by the EB method (1.437), and hence might appear to be "more logical."

In Figure 2 the classical estimate using the more correct Poisson distribution assumption, C2, identifies the same hazardous sites as does the comparable EB method (B2) and recognizes them at an even higher δ than does the EB method. The classical method identifies two sites (7 and 29) that EB does not at $\delta = 9$. Thus, one could rephrase the authors' first sentence in the second paragraph of page 5 in support of the classical estimator as follows: "Criterion B2 (EB) fails to identify as hazardous many sites that are flagged by C1 and C2." The "many" here would be only three sites; however, this is the same number of sites the authors refer to as "many" in their original statement denouncing C1 (Figure 1).

In Figures 3 and 4 we see one site (Figure 3) that the EB method identified and the classical one did not, namely, Site 12, but then in Figure 4, the classical method identifies Site 12 as hazardous when the EB method does not.

In summation, this numerical example does not show, as stated in the abstract, "that some classically based statistical techniques may be prone to err in the direction of false negatives" any more than EB methods. If, after correcting the computational error discovered by Morris, this example still fails to support such a claim, a better example should be found. Otherwise, the would-be user of this methodology is left to conclude that here we have a much more complicated procedure that does no better (maybe even worse) than the more simplistic classical methods. In its present form, this paper appears to do a disservice to a methodology that may, in fact, be superior by (a) containing mathematical errors and (b) presenting a weak example.

DISCUSSION

EMMANUEL QUAYE
Department of Civil Engineering, University of Toronto, Box 10, 321 Bloor Street West, Toronto, Ontario M5S 1S5, Canada.

In this paper the authors suggest ways of estimating the mean and the variance of true accident rates for m sites. The expressions given in the paper for calculating the variance by the method of moments and the formulation of the likelihood function are incorrect.

It can be shown (*1*) that, using the method of moments,

$$\hat{\mathrm{Var}}\{\lambda\} = [1/(m-1)]\left[\sum_{1}^{m} (N_i^2 - N_i)/V_i^2 - (1/m)\left(\sum_{1}^{m} N_i/V_i\right)^2\right] \quad (13)$$

and not

$$[1/(m-1)] \sum_{1}^{m} \left(N_i/V_i - 1/m \sum_{1}^{m} N_i/V_i\right)^2 \quad (14)$$

as given in the paper.

As can be seen from the expressions above, Equation 2 leads to an overestimation of the variance. The difference between the two expressions is

$$[1/(m-1)] \sum_{1}^{m} (N_i/V_i^2) \quad (15)$$

Using data from Tables 1 and 2, the following comparisons, as shown below, can be made. It can be observed that there is a substantial overestimation of the variance by Equation 2.

	Variance	
Data Set	*Equation 1*	*Equation 2*
Table 1	0.0673	0.1410
Table 2	0.1164	0.1759

Similarly, it can be shown (*1*) that the correct likelihood function is

$$\prod_{1}^{m} [\alpha/(\alpha + V_i E\{\lambda\})]^{\alpha}[\Gamma(\alpha + N_i)/(\Gamma(\alpha)N_i!)] \times [V_i E\{\lambda\}/(\alpha + V_i E\{\lambda\})]^{N_i} \quad (16)$$

and not the corresponding expression given in the paper. The performance of this likelihood function was confirmed by simulation (*1*) and shown to give good results.

REFERENCE

1. E. Hauer and P. Garder. Research Into the Validity of the Traffic Conflict Technique. *Accident Analysis and Prevention*, Vol. 18, 1986, pp. 471–481.

AUTHORS' CLOSURE

First, let us state unequivocably that it is truly an honor to have these authors share their thoughts on this paper. It is with pleasure that we respond to their comments.

In our paper, we present a two-stage method for obtaining information about accident rates in a Bayesian fashion. In the first stage, we combine accident histories from various sites within a region to estimate a regional distribution, whereas in the second stage, we use site-specific data to update the regional distribution, thereby obtaining refined estimates of the distribution associated with each site. The comments in the discussions offered by Quaye and Morris concern some of the details associated with the first stage. We shall address their comments first, and save our discussion of Pendleton's criticisms for last.

FIRST-STAGE CONCERNS

Both Quaye and Morris question the manner in which we compute our estimate of the variance of the regional distribution, $f_R(\lambda)$, which has an obvious impact on our initial choice for the parameters α and β. Quaye's objection arises from the fact that we have treated the observed accident rates $\hat{\lambda}_i = N_i/V_i$, $i = 1, \ldots, m$, as our sample of observations instead of the collection of paired values (N_i, V_i). Morris points out that even when the observed rates are used, the manner in which we calculate the sample variance yields a biased estimate of the distributional variance.

Admittedly, because each observed accident rate, $\hat{\lambda}_i$, is derived from two pieces of data, N_i and V_i, Quaye's interpretation of the "sample" may be preferred to ours. However, the method of estimating the variance of the accident rates, presented by Hauer and Garder (*1*), should be considered with caution. These authors verify that the estimate presented by Quaye can provide negative estimates of the variance, a quantity that is necessarily nonnegative. This can cause difficulties, and thus the procedure should be carefully examined before it is used. Although the estimate we used cannot yield a negative sample variance, it does yield a biased estimate of the true variance, as discussed by Morris.

Regardless, herein lies a major difficulty associated with some classical identification techniques. Each of these three methods represents a reasonable or common method used to estimate the variance of the regional distribution, yet each provides a different estimate (although the Quaye and Morris estimates are in very close agreement). Because of the direct dependence of the classical techniques on the computed sample mean and variance, it is clear that the resulting set of sites identified as hazardous depends on the manner in which the data are presented and the statistics are computed. Different estimates of the variance of the accident rates across the region will necessarily lead to different sets of sites that are identified as hazardous. As a result, the tremendous differences in the tabulated values of the computed estimates of the variance presented in the Quaye discussion, corresponding to factors of 2.09 and 1.51 for the first and second data sets, respectively, might cause some concern for the integrity of some of the classical procedures.

To see that the Bayesian technique does not suffer from these same shortcomings, one need only compare the distributions obtained with the two estimation techniques. Clearly, different estimates of the variance of the accident rates will result in different estimates of both the regional and the refined distributions. Figures 5 and 6 show the regional distribution and a representative refined distribution that result from the two estimation techniques. These figures are based on the moments computed from the second data set, because it best shows the differences between the various distributions. The more peaked curve in Figure 5 (i.e., the one with the smaller variance) corresponds to the regional distribution based on the estimate of the variance obtained with Quaye's Equation 13, which is very nearly equal to Morris's estimate. In Figure 6, one can see that the refined distributions are virtually indistinguishable. As we discussed in the paper, this is because in

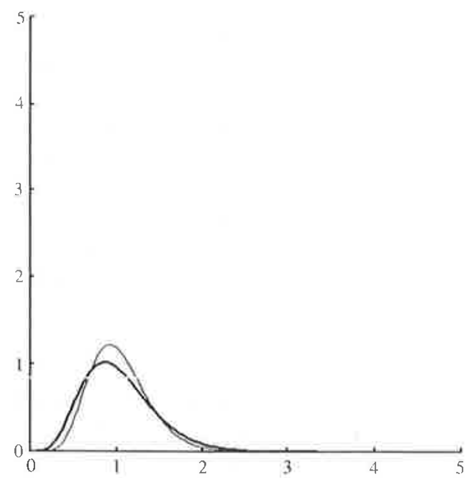

FIGURE 5 Regional distribution (July 1984–June 1986).

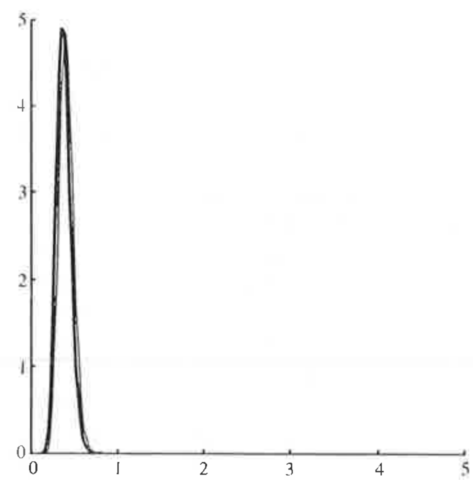

FIGURE 6 Refined distribution (July 1984–June 1986).

updating the parameters associated with the regional distribution to obtain the site-specific distribution, one necessarily overpowers the original parameters (which in our case are based on the computed sample mean and variance) with the site-specific data. Thus, the refined distributions, which provide the basis for the identifications and are therefore of paramount importance, are largely insensitive to the original parameter selection. It follows that one can be reasonably assured that although the classical methods are highly dependent on the sample variance, the Bayesian methods are not.

In addition, both Quaye and Morris question the likelihood function presented in the second section of the paper. In performing our analysis, we worked exclusively with the model that Morris has described as the descriptive model. As such, we have explicitly dealt only with the distribution of the accident rates, which are initially assumed to follow a gamma distribution. This is our Assumption A2. As a result, the likelihood function for the gamma distribution is correct as stated. Naturally, if we had performed our investigation on the basis of N_i, using the inferential model, the negative binomial model suggested by Morris would have been correct, and the likelihood function would change accordingly.

Unfortunately, Quaye's claim regarding the likelihood function (i.e., Quaye's Equation 16) is simply incorrect. A likelihood function is a mathematical entity representing the relative likeliness of the observed data (e.g., $\{(N_i, V_i)\}_{i=1}^{m}$ for a given set of distribution parameters (e.g., α and β for a gamma distribution). Thus, the maximum likelihood parameter estimates for the gamma distribution are those values of α and β that are most likely to have generated the observed data. In a very real sense, they provide the values of α and β that best fit the observed data, although the resulting theoretical mean and variance need not agree with the sample mean and variance. The "likelihood function" offered by Quaye, which also appears in the paper by Hauer and Garder (*1*), is not a true likelihood function. Instead, it is a form of a likelihood function that has been artificially constrained so that the resulting theoretical mean agrees with the observed sample mean. Thus, in general, the parameter estimates that are obtained by using it are not the maximum likelihood estimates, whereas those obtained from the expression in our paper are. In addition, the statement that somehow Quaye's Equation 16 has been "confirmed" by simulation is most disturbing indeed. As a mathematically known function, a likelihood function need not be subjected to empirical validation. Such a validation procedure suggests that there may be a "gray area" associated with the functional definition. Because it is a well-known mathematical entity, there is no such gray area requiring empirical validation.

INTERPRETATIONAL CONCERNS

Pendleton's discussion begins with a simple reiteration of Morris's comments. Because we have already discussed these comments at length, there is no need to further expand on them here. Instead, we shall focus on Pendleton's remaining criticism, which pertains to our interpretation of the results of our empirical study.

First, note that we are not in a position to claim that one technique is superior to another. To do so, one would have to know which sites are actually hazardous so that one can correctly determine which technique tends to provide correct identifications most often. Naturally, in the absence of perfect information, one can only interpret pieces of evidence or results as they become available.

The interpretation of empirical results such as those presented in our paper is necessarily subjective, and solid conclusions are often difficult to reach. Our conjecture that the classical technique, C1, "may be prone to err in the direction of false negatives" is based on an in-depth analysis of the magnitude of the differences in the levels at which sites are identified as hazardous.

To illustrate these differences, consider the three sites from Figure 1 for which techniques B1 and C1 provide differing results, namely, Sites 2, 28, and 29. From Table 1 one can obtain the probabilities computed using the Bayesian technique. A simple algebraic expression identifies the maximum confidence level at which Criterion C1 will identify these sites as hazardous, δ_{max}. Similar quantities can be obtained from the analogous sites associated with Figure 3 (i.e., Sites 12 and 29). These values are summarized in Table 6.

On the basis of this information, it seems clear that B1 provides very strong evidence that Sites 4 and 28 are hazardous (0.9813 and 0.9891, respectively), whereas C1 provides substantially weaker evidence (0.8874 and 0.9037, respectively). Similarly, although Site 12 receives a lower degree of support from B1 than do Sites 4 and 28, it still receives a substantially higher level of support from B1 (0.9083) than from C1 (0.7753). Of course, in the first data set, Site 29 receives a lower level of support from B1 (0.8908) than from C1 (0.9512), but the difference is smaller in this case. Because we believe that this type of analysis provides a better understanding of the difference between the methods than does Pendleton's method, we stand by our earlier claim. Of course, further investigation of the differences between the Bayesian and classical meth-

TABLE 6 SUMMARY OF DIFFERENCES BETWEEN B1 AND C1

Figure 1:	site	prob. (B1)	δ_{max} (C1)
	4	0.9813	0.8874
	28	0.9891	0.9037
	29	0.8908	0.9512
Figure 3:	**site**	**prob. (B1)**	**δ_{max} (C1)**
	12	0.9083	0.7753
	29	0.9951	0.9812

ods is called for. In addition, because Pendleton's observation regarding the apparent reordering of Sites 4 and 29 in Figure 1 is eloquently explained in Morris's discussion, we shall not endeavor to expand on his explanation here.

In conducting the research reported in this paper, it was our intention to offer a Bayesian technique for identifying hazardous intersections and to begin to understand how our technique differs from some of the classical techniques. It was not, as Pendleton states, to show that the Bayesian methods "are superior to two classical methods." Pendleton suggests that because our data set fails to provide evidence of superiority of the Bayesian method, another example should be "found" that will support such a claim. The purpose of this research was to explore the truths of a situation, not to discard or manufacture data in an effort to support a desired result.

CONCLUSION

In conclusion, we would like to agree with the closing remarks made by Morris. The application of Bayesian analyses to accident data does appear to provide a fruitful avenue of exploration. There are numerous modeling techniques to be explored. In addition, a further understanding of the differences in the results provided by Bayesian and classical identification techniques is of obvious importance. The pursuit of knowledge in this exciting field promises to offer its own rewards.