TRANSPORTATION RESEARCH
CIRCULAR

Number E-C179 December 2013

Theory, Explanation, and Prediction in Road Safety

Promising Directions

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH CIRCULAR E-C179

Theory, Explanation, and Prediction in Road Safety

Promising Directions

Prepared by
James Bonneson and John Ivan

for the
Future Directions Subcommittee
Task Force for the Development of a Highway Safety Manual
Transportation Research Board

November 2013

Transportation Research Board
500 Fifth Street, NW
Washington, D.C.
www.TRB.org

TRANSPORTATION RESEARCH CIRCULAR E-C179

Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001
www.TRB.org

Javy Awan, Production Editor; Jennifer Correro, Proofreading and Layout

Preface

On November 20 and 21, 2008, the FHWA and TRB cosponsored a workshop titled, Future Directions in Highway Crash Data Modeling. FHWA's Office of Safety Research and Development and TRB's Technical Activities Division were responsible for organizing the workshop. They enlisted the help of the Future Directions subcommittee of the TRB Task Force for the Development of a Highway Safety Manual. The subcommittee provided technical input on the workshop agenda and prepared this document. Primary authors from the subcommittee include James Bonneson, Texas Transportation Institute, and John Ivan, University of Connecticut.

The objectives of the workshop were (*a*) to explore promising future directions in highway crash data modeling and (*b*) to identify potential areas for advanced research to provide a theoretic foundation for explaining crash causation. The focus of the workshop was on the effect of highway infrastructure elements on safety. Goals of the research identified by workshop participants are the further development of science-based safety evaluation and the development of more stable, reliable, and transferrable highway safety predictive models.

Twenty-five invited traffic safety researchers and practitioners attended the workshop. Experts in highway safety data analysis were brought together with experts from allied fields (such as epidemiology, public health, and human factors) to share their experiences and methods for quantifying treatment effects. The workshop agenda is provided in Appendix A.

Through breakout groups and plenary sessions, the participants were asked to describe critical issues and challenges; explore alternative modeling approaches and concepts; and identify promising new directions for explaining the contributing causes of crashes. This document summarizes the discussion and describes potential areas of advanced safety research identified by the participants.

Acknowledgments

Sincere thanks to all of the individuals who participated in the workshop. Special thanks go to John Mason, Associate Provost and Vice President for Research, Auburn University, for facilitating the discussion in the plenary sessions and organizing the feedback received from the breakout sessions. Thanks go also to Ray Krammes, FHWA, and Richard Pain, TRB, for organizing the meeting, making local arrangements, and providing coordination with other agencies and committees.

The contributions of all who participated in the workshop are gratefully acknowledged. The discussions were spirited and the information provided was invaluable. Their subsequent comments on the draft of this document are also appreciated. The workshop participants are identified in Appendix B. Of particular note are the thoughtful contributions of Ezra Hauer and Gary Davis, which are provided in Appendices C, D, and E.

Contents

Introduction

An important role of highway safety research is in the development of information to be used by agencies designing, operating, and maintaining the nation's streets and highways. Information that can be used to make the roads safer and crashes less severe is of particular interest to these agencies.

The engineers who use safety information are responsible for the decisions made from it, and are obliged to make some judgment about its credibility and applicability. They base judgments of credibility on the degree to which the information is consistent with their experience, the experience of others, and similar information from other sources. Information that is consistent in this manner is considered to be credible. Judgments about the applicability of the information relate to whether the information is believed to be accurate when applied to a particular location with specific traits. In some instances, the information is qualified to apply only to specific conditions. If these conditions are not present at the locations of interest to the practitioner, then the information is not transferrable and is considered to have limited applicability.

The information that is the subject of discussion in this document is considered to be quantitative. That is, it can be used to quantify either the safety of a highway facility or the change in safety associated with a change in the condition of the facility. The safety of a facility is defined as the number of crashes, by kind and severity, expected to occur on the facility during a specified time period (*1*).

Highway safety research over the past several decades has developed qualitative and quantitative information that has been effectively used by highway agencies to reduce the traffic fatality rate and to generally make the nation's highways safer. Much of this research has focused on developing some of the more obvious and widely applicable strategies (e.g., increase seat belt usage, improve roadside safety hardware, improve vehicle crashworthiness, etc.). Information about new strategies and treatments may not be as obvious, or as applicable to all facilities. For this reason, more sophisticated scientific methods may be needed to quantify the effectiveness of new strategies. An important question going forward is how best to use the limited available research funding to achieve further improvements in highway safety (*2*).

Highway safety research over the past several decades has relied on the use of statistical methods to quantify the safety of a facility and to estimate the relationship between a change in facility condition and a change in facility safety. These methods have provided essential insight regarding conditions that are clearly unsafe. However, there are also some challenges associated with sole reliance on statistical methods in terms of their limited ability to deal with counterfactual prediction, driver adaptation, unobserved heterogeneity, and data quality. These challenges may make it difficult to achieve further advancement using statistical methods alone. They may partially explain the smaller reduction in fatality rates in the early 2000s (relative to, for example, the 1980s) (*3*).

The objectives of this report are to (*a*) review scientific methods available to develop highway safety information; (*b*) highlight the challenges related to the use of the methods; and (*c*) describe potential areas and components that could help advanced safety research to overcome these challenges.

The methods of interest include those used for study design, modeling, and statistical analysis. Research is identified that would contribute to a theoretic foundation for explaining cause-and-effect relationships that occur in the driving environment: relationships that can ultimately be integrated into safety prediction models that are robust, transferrable, and scalable.

The focus of the discussion here is on the research needed to develop new methods (and the refinement of existing methods) for either quantifying the safety of a facility, or estimating the relationship between a change in facility condition and a change in facility safety. In this regard, methods used to develop highway safety information based on surrogate measures appear to be a fruitful area of investigation, especially when the linkage between a surrogate and safety is established.

The focus of this circular is on information related to the effect of highway infrastructure design and operation on traffic safety. However, it is recognized that the understanding of safety effect will also require an understanding of driver behavior, vehicle performance, and crashworthiness as it relates to crash occurrence and crash severity.

The discussion is focused on the use of crash data as the basis for developing safety information. However, it is recognized that naturalistic driving systems and driving simulators represent complementary sources for safety information. This information may be particularly helpful in the development and testing of theories that predict causality based on driver behavior.

The remainder of this circular consists of two main parts. The first part provides some background information on the various methods that are available to develop quantitative safety information. The second part describes potential advanced safety research, based on the input of individual participants in the November 2008 workshop.

Background

This part of the circular provides some background discussion on methods used to develop highway safety information. The first section describes the various methods, the challenges faced when using them, and the potential for their improvement. The second section describes potential research and development needs to improve methods for highway safety evaluation. The last section provides a summary of the issues and underscores the need for further advancement in the science of highway safety modeling and statistical analysis.

METHODS FOR ESTIMATING SAFETY AND SAFETY EFFECT

This section discusses various methods used to develop highway safety information. This information may take the form of a model for estimating the safety of a highway facility. It may also take the form of a model for estimating the change in safety associated with a change in facility condition (e.g., an intervention or treatment). The model's functional form can range in complexity from a single constant, to a multivariate function, to a series of nonlinear equations that are used in sequence. In theory, a model developed to estimate the safety of a facility may also be used to estimate safety effect by successive application of the model, once with and once without the change in condition.

Model Analysis Scale

Models developed for estimating safety (and those for estimating safety effect) can be described in terms of their analysis scale. The scale is described as microscopic when the basic unit of evaluation is the individual vehicle at a facility and the response variable is the probability of a crash (or similar surrogate) for a given trip-based time interval. In contrast, the scale is described as macroscopic when the basic unit of evaluation is the facility and the response variable is the expected crash frequency (or similar surrogate) for the population of vehicles, as averaged over a long time interval. The length of time represented by the averaging interval for the macroscopic scale is often dictated by sample size requirements and is inherently longer than the trip-based time interval.

In application, a microscopic model can be used to estimate the expected crash frequency (or safety surrogate) for the population of vehicles at a given facility. This estimate is obtained by evaluating the safety of a distribution of vehicle trips over a long time interval and aggregating the responses. An attractive feature of a microscopic analysis is that it allows the safety of each individual vehicle to be evaluated for the environmental conditions present at the trip time. In contrast, for macroscopic analyses, the facility is evaluated for one set of environmental conditions that are considered representative over the analysis period and for which one response variable is obtained that represents an average for that period.

One challenge with developing models is that of data aggregation (*4*). The aggregation of data can lead to bias in model calibration coefficients and inaccurate inferences about safety effect. Aggregation of data can allow certain facility traits to have undue influence on the model structure or mask safety effect at individual facilities. This problem is particularly troublesome when developing highway safety models given the infrequent nature of crashes and the limitations of crash and infrastructure data. Safety surrogate measures occur more frequently than crashes and desired sample sizes are obtained in shorter time periods, thus minimizing some of the data aggregation issues. However, their use presumes a foundation has been established that links the surrogate measure to safety.

Researchers desiring to develop transferrable models often seek to expand the spatial extent of the analysis by evaluating additional facilities and pooling the data. This option is also attractive from the standpoint of increasing sample size. However, the problems of data aggregation are magnified, especially if one or more facility attributes is unmeasured (i.e., unobserved heterogeneity is increased).

Modeling Approach

Models developed for estimating safety, or safety effect, can also be described in terms of modeling approach. The modeling approach can be generally described as "statistical" or "structural." The traditional statistical modeling approach uses the assembled data to guide the development of model functional form and to quantify correlation. However, as Hauer (*5*) notes, modelers have no theory to guide their choice of model form nor recognized procedures for formulating this form. Current practice is to pose a few alternative functions and choose the one that fits the existing data best. As a result, model functions tend to have overly simplistic structures and more empirical constants than variables. Improperly specified functional forms can be as problematic as having correlated independent variables or unobserved heterogeneity in the data. Moreover, an improper functional form can induce nonexistent correlations and artificially increase the error variance.

Hauer (*6*) points out that statistical model development could be improved if prior knowledge about specific functional forms and factor effects was used by subsequent modelers in their development of statistical models. Hauer suggests that the key to obtaining credible models is the consistency with which the same effect is found in data, and the key to scientific advancement is the degree to which these findings are then used in subsequent model developments. However, current practice in statistical model development is to use only the assembled data to guide model development with the result all too often being a different function and (a different set of model coefficients) for each research effort that investigates a similar factor or treatment.

The structural modeling approach (also referred to as mechanism-based modeling) is characterized as using background knowledge to describe the relationship between input (exogenous) variables and the response variable. The description is in the form of equations that describe causal dependencies between model variables (*7*). The equations can be based on theoretic constructs or empirically derived relationships. They describe vehicle–driver system in terms of physics, driver behavior, and the interaction of vehicles with the highway environment. More detail on structural modeling approaches is provided in Appendix D.

From a mathematical modeling standpoint, the structural modeling approach can be characterized as clear (or white) box modeling because prior information is used to develop the model formulation. The advantage of a clear-box model is that it provides a framework for scientific advancement and a transparency to the interpretation of causal dependencies. In contrast, traditional statistical modeling approach can be characterized as closer to black box modeling because there is limited reliance on prior information to develop the functional form and no element of causality (only correlation).

Model Calibration Method

This section describes the various study designs used for model calibration. The focus of the discussion in this section is based on the prediction of expected crash frequency for a facility; however, other safety measures or surrogates could apply.

One criterion for establishing the study design is based on the intended model application. If the model is intended to quantify the safety of a facility, then recent crash history data are considered the best representation. Time-series data can be used to expand the temporal extent of the analysis. If the model is intended to be transferrable to other facilities, then cross-section or panel data are often used.

If the model is intended to quantify the change in safety associated with a change in condition at a facility, then before-after data are often used. If the change in safety is intended to be transferrable to other facilities, then before-after data from multiple facilities are often used.

When data for multiple facilities are used, statistical tests for unobserved heterogeneity can be used to determine if there are unexplained systematic differences in safety (or safety effect) among facilities. Ideally, the model function is refined to account for these differences. When the heterogeneity cannot be accounted for in this manner, then it is accounted for in the model variance structure and coefficient confidence intervals. In this regard, the negative binomial structure with a gamma-distributed error term is often used to account for unobserved heterogeneity. However, Anastasopoulos and Mannering (*8*) argue that a random-parameters analysis is more effective at accounting for unobserved heterogeneity across road segments. In contrast, Shugan (*9*) argues that unobserved heterogeneity should be considered to be random error when testing theories.

The remainder of this subsection briefly summarizes the challenges associated with the use of before–after and cross-section data for safety model calibration. These are the most common types of data used for this purpose. However, it is useful to note here that there other possible model calibration approaches. Of particular interest is the method proposed by Davis (*4*) for estimating the effect of a change in safety associated with a change in condition at a specific facility. This method is based on structural modeling of individual crash events at a microscopic scale. Davis acknowledges that this approach may yield biased estimates of safety effect due selection bias. However, the promise of this approach is that it describes a plausible causal chain that explains the effect of the change in condition. Davis poses that this approach provides essential insight in the search for causation, identification of confounding factors, and the formulation of effective intervention strategies.

Before–After Study

For highway safety applications where the data are observational, the before–after study is generally recognized as the most appropriate method for quantifying the change in safety due to a change in facility condition. However, as noted by Hauer (*6*), before–after studies are not practically feasible for the evaluation of many specific types of change. He states that "…opportunities to do observational before–after studies about, say, the safety effect of change in horizontal curvature, road grade, lane width, median slope etc. are few and imperfect. This is so, partly, because when a road is rebuilt usually several of its attributes are changed at once and it is difficult to assign the result to any single causal factor. In addition, the rebuilding of a road often changes it to such an extent that it may not be regarded the same unit after reconstruction" (*6*).

An additional concern about the before–after study is that it focuses, intentionally, on one change in condition. The before–after study can be seen as limiting when the changed condition represents an entity that can vary in dimension (e.g., curve radius, yellow interval duration). Effectively multiple studies would need to be conducted to calibrate a model for estimating safety effect over the full range of dimensional values. The problem increases in exponential magnitude when the safety effect is a function of multiple variables (e.g., number of lanes, speed).

Elvik (*10*) describes a procedure for calibrating a model of safety effect based on the results of multiple before–after studies. He refers to this type of model as an "accident modification function." He demonstrated the procedure for two example applications and discussed the challenges associated with the use of this procedure.

Cross-Section Study

The cross-section study is often used to quantify the safety of a facility. The model typically developed from cross-section data is referred to as a safety performance function. This model is widely recognized as an effective means of summarizing the safety of a facility represented in the calibration data set, as may vary over time. However, it is often used to describe the safety of a similar facility not represented in the data set. This application has some risk and local calibration of this type of model is generally performed to minimize related problems.

The opportunity to use cross-section data to quantify safety effect has been explored in depth by Hauer (*6*). When used to quantify safety effect, a calibrated model is used twice—once to estimate the safety of a specific condition and then again to estimate the safety with a different condition. The difference in safety estimates is then inferred to represent the effect of the change in condition on safety. This application is limited to models that contain a variable that associates with the condition being changed (e.g., lane width). The assumption is that the model accurately predicts the change in safety that would be incurred by those facilities with condition "c" should the condition be changed to "t," all other factors being the same. However, if the facilities with condition "c" and those with condition "t" are not homogeneous, then the change in safety cannot be fully attributed to the change in condition (i.e., the estimated effect is considered to be confounded).

Hauer (*6*) explored the extent to which potential confounding can be removed or corrected, thus allowing the use of cross-section studies to quantify safety effect. He examined the use of attribute restriction, attribute matching, and attribute equalization (through parametric regression). The first two methods were identified as being more effective correction methods; however, they have some limitations that preclude their use in all situations. Hauer (*6*) identified several challenges that must be overcome before attribute equalization is a viable correction method. These challenges are

- The functional relationship chosen for the regression model can influence the result;
- A tendency for modelers to rely only on the data to quantify model coefficients (i.e., in ignorance of prior knowledge); and
- Statistical modeling techniques currently used are not sufficiently structured that different researchers will converge on similar model forms and coefficient values.

He notes that Bayesian regression incorporating informative priors (based on accumulated knowledge) may overcome some of these challenges.

RESEARCH AND DEVELOPMENT NEEDS

The discussion in the preceding section suggests that advancements in both the science of modeling approach and the science of model calibration are needed. This need is highlighted in a recent TRB Special Report. This report states that "crashes are complex events resulting from a combination of factors affecting the driver, the vehicle, and the roadway. Hence, understanding the key factors resulting in a crash and developing effective countermeasures require a rigorous science-based approach that seeks to identify and isolate many contributing and often interrelated effects" (*2*).

Evidence

The content of the *Highway Safety Manual* (HSM) provides some useful perspective on the types of safety information needed by practitioners for road safety management (*11*). Among the many types of information provided, there are (*a*) models that can be used to quantify the safety of a facility and (*b*) models that can be used to estimate the relationship between a change in facility condition and a change in facility safety. The methods used to develop these models were largely focused on finding statistical association.

The development of the HSM caps five decades of safety research. However, progress is seen as limited by some in terms of the development of credible, transferrable knowledge. For example, a review of safety research in the past decade indicates that little of what has been learned is used in subsequent research to advance the science of safety estimation (especially as it relates to identifying underlying mechanisms related to crash causation). Hauer (*6*) suggests that individuals who have attempted a critical review of the safety literature "...will attest to the fact that many of the research reports found will be quickly discarded. They will be thought too deficient in method, too small to draw conclusions from, inconclusive, obsolete, of obscure message, biased, or otherwise seriously flawed. In the end, one is left with very few studies that are not obviously unreliable and the results of which do not contradict each other" (*6*).

A review of the safety literature by Davis (*4*) revealed concerns of a similar nature. As an example, he cited research on the effect of an increase in speed (notably, an increase in the national maximum speed limit). He identified two reviews of the literature on speed and safety that were commissioned by the TRB (*12*, *13*). The two independent reviewers cited 73 and 65 sources, respectively, before concluding that "…although evidence tended to support the notion that accident risk increased with speed, more study was needed to determine when changes in speed limit affect accidents or to predict the sizes of these effects" (*4*). In the highway safety arena, this large number of sources addressing a common topic is rare. It is an indication that the research topic represents an area of significant information need and corresponding research investment. The fact that so little is known after this level of resource expenditure could be taken as an indication of the need for further improvement in the science of safety estimation (hopefully, this result will not be construed by research funding agencies that the question cannot be answered).

Establishing a Theoretic Basis for Explanation

Elvik (*14*) observed that "One of the major problems of road safety evaluation research is the fact that most of this research does not have a strong theoretical basis, which guides the design of studies and the interpretation of study findings. The lack of a strong theoretical basis for research means that few results of road safety evaluation studies can be ruled out on theoretical grounds" (*14*). He suggests that a road safety treatment normally influences road safety by way of two causal chains: (*1*) the engineering effect and (*2*) the human behavioral feedback to engineering changes (i.e., the behavioral effect). He offered that the engineering effect of a treatment can be explained in terms of one or more of nine engineering risk factors. These factors include:

- Kinetic energy,
- Friction,
- Visibility,
- Compatibility,
- Complexity,
- Predictability,
- Individual rationality,
- Individual vulnerability, and
- System forgiveness.

Elvik (*14*) posed that the engineering effect of a treatment is described by the change in one or more of the risk factors, relative to a threshold value associated with perceived unsafe condition. This relationship he called the structural margin of safety. Examples of variables used to quantify the structural margin of safety included sight distance, lateral distance to roadside hazard, speed, friction demand, etc.

The behavioral effect of a treatment reflects driver adaptation to changes in facility design, operation, or level of enforcement. These changes influence the driver's perception of safety and, thus, their response to the change. Elvik (*14*) identifies several factors as influencing road user behavioral adaptation, they are

- How easily a measure is noticed;
- Antecedent behavioral adaptation to basic risk factors (that reduces the potential for further adaptation due to a specific change);
- Size of the engineering effect on generic risk factors;
- Whether or not a measure primarily reduces injury severity;
- The likely size of the material damage incurred in an accident; and
- Whether or not additional utility can be gained.

The engineering effect and the behavioral effect can be offsetting to varying degrees, depending on the ease with which the change is detected by drivers and the driver's perceived consequences of a crash.

The development of theoretic principles to guide safety evaluation and to form a basis for model development will not be without challenge. Some previous attempts at theoretic explanation have not proved fruitful and, in fact, may have been misleading. Progress in the development of safety-based theories may take longer than for other areas (e.g., traffic flow theory) and will require a thorough understanding of driver behavior and adaptation. Credible advancement will require that all theories are founded in empiric evidence. The tests needed to confirm theories should be more rigorous than those used to establish statistical correlation (*9*). More discussion on these issues and challenges is provided in Appendix C.

Structural Modeling Approach

Other scientific fields have experienced similar challenges to the advancement of methods for scientific evaluation. Highway safety researchers can learn from these allied fields. For example, consider Paul Thagard's (*15*) three-stage model about progress in explaining disease. In the first stage, researchers identify reliable associations between disease occurrence and other factors. Then, in the second stage, the associations are used to establish causal connections. Finally, in the third stage, researchers work to identify the underlying disease mechanisms. When the mechanism has been identified, the range of possible treatments increases. But while statistical methods are usually central in the first two stages, different methodologies are often needed for the last stage. If progress does not progress through these stages, then scientific advancement will likely be constrained and research investment will be inefficient.

The rational management of road safety requires causal models that allow one to predict the consequences of design and operational decisions (*16, 17*). Theory is used to dictate the direction of possible causality in structural models. Data is necessary to confirm the theory. The theory is then used to disentangle confounding influences and guide study design.

Davis (*4*) argues that structural modeling is the appropriate method for explaining causal relationships in highway safety. A structural model would consist of deterministic mechanisms that draw on background knowledge concerning how the driver–vehicle–road system behaves. Using a structural model, prediction of the aggregate causal effect of a specified treatment becomes a bottom-up approach. First, the relevant mechanisms for a specific type of crash are identified. Then, they are used to quantify the causal effect of the treatment on each mechanism. Finally, the frequencies of the mechanisms are aggregated for the facility of interest. A cited benefit of using a structural modeling approach is that prior knowledge of underlying mechanisms or relationships helps guide experimental design, data analysis, and interpretation. More information on the structural modeling approach is provided in Appendix D.

Multidisciplinary Research

A recent TRB publication that examined the practice of highway safety research cited a need for increased use of multidisciplinary teams on individual research projects (*2*). The reason for this need was an observation that crash events represent the outcome of a complex combination of driver, vehicle, and environmental conditions that combine in a unique and interrelated manner. A fuller understanding of these events, conditions, and interrelationships would be realized when the research teams include individuals with knowledge in a variety of discipline areas, including: engineering, psychology, human factors, economics, statistics, education, law enforcement, systems analysis, marketing, biomechanics, and public health.

SUMMARY

In general, what is known on a scientific basis begins with the observation of events, examination of data, and reflection on current knowledge. Questions naturally arise in the form of a desire for fundamental understanding or a need to improve decision making. Explanations and hypotheses are proffered (the alternative is to accept the status quo). The explanation posed is used to guide experimental design and the postulation of explanatory relationships. Data are gathered and the hypothesis tested. The findings are promulgated and there is professional discourse on the topic. Knowledge is gained. The cycle repeats using the new knowledge as a starting point. With each cycle, the understanding grows deeper, the findings become more consistent, and the theoretic explanation evolves by becoming more refined and credible.

In engineering disciplines, statistical methods have provided an essential tool for formulating an initial, preliminary understanding of the relationship between variables. However, an essential next step in the evolution of understanding is the development of a quantitative, theoretical basis for explaining the correlations uncovered by statistical analysis. This theory then becomes the foundation upon which the understanding of cause and effect is based. The burden placed on statistics and data is lessened. Resources are shifted to the development of structural models with a solid theoretic foundation, where data are used as a means of confirming the theory and calibrating the model. Future generations of professionals are educated on the theoretic principles and are equipped to contribute to this body of knowledge upon entry to the workforce. Practitioners are provided the calibrated models as a basis for making informed decisions.

In highway safety evaluation, much is known about the relationship between variables. However, little is known from a quantitative, theoretic perspective about the complex processes that combine to trigger a crash. Further advancement requires a more rapid development of theoretic principles. Initially, these principles may be simple and address only the main effects. Statistics will be needed to confirm the theoretic principles. Theoretic relationships would then be used to construct structural models that explain the connection between input state, intermediate variables, and outcome. These models will require calibration and validation using data. The debate in the profession will be on the rigor of the theoretic principles used in, and the robustness of, the models. In some cases, established theory will allow subsequent researchers to fill the void of missing variables (or hard to measure data) and minimize problems associated with data quality. In time, the theoretic principles will be refined and new ones derived; more robust models will be developed. Ultimately, the theoretic principles will be sufficiently mature that cause and effect will be understood and accepted, and a wide array of safety prediction models will be available to practitioners for decision support.

In summary, the field of highway safety engineering has advanced to a point where it could transition to a more formal development of theoretic principles to ensure the effective use of limited research funding and to establish a foundation for future scientific advancement. Moreover, the development of theoretic principles will provide a sound basis for educating new safety professionals and advancing safety science. These principles would form the basis for the development of models for (*a*) estimating the safety of a facility for a given set of conditions and (*b*) estimating the change in safety as a result of a change in condition. The models would form the basis for future editions of the HSM and the safety simulation models that might complement it.

Research Themes

On November 20 and 21, 2008, FHWA and TRB cosponsored a workshop, Future Directions in Highway Crash Data Modeling. The objectives of the workshop were to explore promising future directions in highway crash data modeling and to identify potential advanced safety research that could provide a theoretic foundation for explaining crash causation. Through breakout groups and plenary sessions, the participants were asked to describe critical issues and challenges, explore alternative modeling approaches and concepts, and identify the most promising new directions for explaining the contributing causes of crashes. This part of the document summarizes the discussion and describes the areas and components of potential advanced safety research identified by the individual participants in the workshop.

Four major research theme areas were identified, as shown in Table 1. Each theme area is shown in the table to consist of several research components. These components identify specific research projects to address the issues associated with each theme area. These areas and components are described in more detail in the remaining sections.

THEME AREA 1: STRUCTURAL MODELING

Statistical forecasting is a powerful methodology, but by its nature cannot be used to prove causality. Structural modeling has been proposed as an approach to crash analysis that can identify how and why crashes occur. This section describes research to advance the

TABLE 1 Research Theme Areas and Components

Theme Area	Research Component
Structural modeling	Document state of the art in structural modeling. Development of driver behavior models for structural modeling. Research in the area of structural modeling. Relating statistical and structural models using stochastic simulation.
Surrogate measures of safety	Surrogate scoping effort: definition, criteria, needs, and priorities. Evaluate and validate candidate surrogate measures. Demonstrate application of most promising surrogates in appropriate situations. Apply surrogates to develop a fundamental understanding of safety issues. Derivation of crash opportunity measures.
Improving the evaluation of intervention effectiveness	Taxonomy of crash prediction models: strengths, weaknesses, applications. Technical improvements to the empirical Bayes approach. Investigation of the use of randomized trials for road safety research. Analysis methods for evaluation of top-performing entities. Framework for developing and testing the adequacy of alternative models or modeling methods.
Multiscale safety and driver behavior modeling approaches	Driving simulator and roadway data complementarities. Models that capture micro–macro transition (multiscale modeling). Road user adaptation in multiple time scales (milliseconds to months).

understanding and application of the structural modeling approach in the area of highway safety evaluation.

Document State of the Art in Structural Modeling

This research component would document the state-of-the-art in areas related to structural modeling. Topics of investigation include: methods for aggregating safety information, margin of safety analysis, vehicle crash mechanics, road–user behavior models, safety surrogates, and statistical methods for model calibration. This research would provide a knowledge base for subsequent research projects.

Development of Driver Behavior Models for Structural Modeling

Develop realistic representations of road–user behavior for use in structural models, with consideration given to all road users, including drivers, pedestrians, and motorcyclists. The focus would be on driver behaviors that are associated with engineering inputs. Challenges to advancement in this area are recognized to be significant. Demographic variations in driver behavior will be a major challenge.

Research in the Area of Structural Modeling

This research would develop a framework and procedures for constructing structural models for highway safety prediction at the microscopic level of detail. The research would develop techniques for constructing a graphical representation of the driver–vehicle–environment system and then use this representation to define mechanism models and empirical models as part of causal chain.

The knowledge and skills to construct comprehensive structural models are likely to include expertise in several disciplines that are not usually covered in civil engineering graduate education. Therefore, this research component would include a series of educational opportunities and workshops that demonstrate the benefit of structural modeling and describe the procedures for constructing them.

Relating Statistical and Structural Models using Stochastic Simulation

Structural models show promise for accurately describing how and why crashes occur at a microscopic level of detail. On the other hand, macroscopic statistical models are able to describe variation in crash occurrence on the basis of large samples that are not practical for structural models. This research would explore the use of stochastic simulation models as a means of bridging the gap between the structural and statistical modeling methods, and would also develop methods for using microscopic structural models, simulation, or both for the development of macroscopic safety prediction models.

THEME AREA 2: SURROGATE MEASURES OF SAFETY

Surrogate safety measures provide an important means of evaluating alternative safety treatments, especially when these treatments do not lend themselves to evaluation with crash data. Surrogates allow engineers to assess facility safety in a shorter time period and at a lower cost than would otherwise be required to when using crash data. Evaluation of surrogate measures could also provide insight into the crash causation process. This section describes potential research to ensure the maximum utility of surrogate safety measures.

Surrogate Scoping Effort: Definition, Criteria, Needs, and Priorities

An initial research component would focus on a state-of-the-art review of knowledge in the area of surrogate safety measures. This effort would identify and define candidate surrogate measures. It would establish criteria to assess the validity of alternative surrogate measures. The criteria would consider the use and usefulness of each measure in road safety evaluation. Thereafter, research would identify potential roles for each candidate measure (e.g., for countermeasure evaluation, or as an independent variable in a safety prediction model).

Evaluate and Validate Candidate Surrogate Measures

This research component would consist of a series of separate research projects. Each project would evaluate and validate one candidate surrogate safety measure or one specific class of related measures (e.g., surrogates from simulation modeling, surrogates from field studies). A key element of this evaluation is to quantify the strength of the relationship between the surrogate measure and crash frequency. The evaluation would also be used to identify the potential applications and best use of each surrogate measure. For example, is the measure appropriate for the evaluation of specific roadway types, vehicle types, and geographic areas? Alternative methods for measuring a surrogate in the field would be evaluated and described. The level of effort to collect each surrogate measure in the field would also be addressed.

Demonstrate Application of Most Promising Surrogates in Appropriate Situations

This research component would consist of a series of separate research projects. Each project would demonstrate the application of one surrogate safety measure (or class of related surrogate measures). The measures considered for this activity would be those evaluated in the previous research components and identified as having the most promise. It is envisioned that this research component would include a feedback loop that would further refine the surrogate measure definition, role, application, data collection method, and utility.

Apply Surrogates to Develop a Fundamental Understanding of Safety Issues

This research component would consist of a series of separate research projects. Each project would evaluate the ability of each safety surrogate (or class of related surrogate measures) to describe facility safety, or to quantify the safety effect of a treatment. Also evaluated would be the potential use of surrogates in structural models and safety prediction models.

Derivation of Crash Opportunity Measures

The objective of this research component is to evaluate the suitability of incorporating the concept of crash opportunities or vehicle interactions into crash prediction modeling. These types of exposure metrics would replace the traffic volume variable currently used in most macroscopic models. The idea is to find a metric that is (*1*) more consistently correlated with the crash count than the traffic volume and (*2*) easily observed or can be accurately estimated from traffic volume. The research would focus on the development of exposure metrics for specific crash types, as may occur at specific types of facilities.

THEME AREA 3: IMPROVING THE EVALUATION OF INTERVENTION EFFECTIVENESS

Understanding the effectiveness of specific road safety interventions continues to be a critical issue in road safety. Advances have been made in many methods of analysis, but there remains room for improvement in their application to highway safety evaluation. This theme area is intended to focus research on the improvement of statistical methods. The objective is to increase the robustness of existing methods as well as investigate the viability of new methods.

Taxonomy of Crash Prediction Models: Strengths, Weaknesses, and Applications

This research component would develop a synthesis document that catalogs the various forms of models that are used for predicting crash frequency. It would also describe the theoretical or practical basis for each model form along with its strengths and weaknesses. The information obtained through this research would likely be of value to researchers focused on other components of this theme area. Therefore, this research component would precede the other components of this theme area.

Technical Improvements to the Empirical Bayes Approach

The strength of the empirical Bayes approach for evaluating safety effect is in the fact that it compares the observed safety condition after the treatment with the safety condition expected in the time period after the treatment if the treatment had not been applied (i.e., what would have happened without treatment, also referred to as the counterfactual prediction). It has been noted however, that current methods for counterfactual prediction do not accurately account for the effect of time trends between the before and after periods. It has also been suggested that further research is needed into the specification of the error term used in estimating the counterfactual outcome. More generally, this research would develop better methods for counterfactual prediction.

Investigation of the Use of Randomized Trials for Road Safety Research

Medical, pharmaceutical, and public health researchers have used randomized trials in their investigations of the effectiveness of medical and pharmaceutical treatments for decades. In highway safety research, it has commonly been asserted that this is not possible because of the physical and institutional circumstances inherent to road operations. This research component would identify opportunities to design controlled experiments to quantify the effectiveness of specific road safety treatments. These experiments would meet the objectives of a randomized trial framework within the physical and institutional constraints inherent to road operations. The identified opportunities would be prioritized and the experiments conducted.

The findings from this research component would be used to document the potential promise of randomized trials for highway safety evaluation and the treatments to which it may be most amenable. If possible, the findings would be compared to those obtained from more traditional before–after study findings, opportunities for improvement to the before–after study method identified such that it would be more likely to yield results that are as accurate as those from a randomized trial.

Analysis Methods for Evaluation of Top-Performing Entities

In road safety research, most of the focus has been on the most dangerous road sections or intersections (i.e., the worst performing entities) to discover what needs to be done to make them safer. In contrast, sports researchers frequently focus on the top-performing athletes to find out what they are doing to facilitate their superior performance. It has been suggested that road safety researchers could focus on the top-performing roads and intersections, that is, those with few or no crashes, to see what can be learned about making all roads safer. This research approach would develop and implement a plan for investigating the merit of this approach. It would document the findings and describe procedures for using this method to quantify the safety of a facility or the effect of a treatment on safety.

Framework for Developing and Testing the Adequacy of Alternative Models or Modeling Methods

New and improved methods for the analysis of data will continue to be developed. What is missing is a tool for assessing the extent to which a method succeeds in identifying the cause-effect relationship in the data. The development of such a tool is feasible. It could consist of constructing artificial realistic data (ARD) wherein the underlying cause–effect structure is known to an independent agency. The data would then be used by researchers who have no knowledge of the cause–effect structure and who would apply the method they wish to test. The degree to which they succeed would then be evident. A more complete discussion on the role of ARD in road safety research is provided in Appendix E.

THEME AREA 4: MULTISCALE SAFETY AND DRIVER BEHAVIOR MODELING APPROACHES

The most critical, and perhaps least understood, aspect of road safety is the human component. Human error is an often cited crash contributor, yet very little is known about the actual behaviors leading to the crash. A better understanding of driver actions (and reactions) in response to the driving environment would provide insight into the selection and design of safety treatments. Coupled with this investigation is research into driver risk assessment and its influence on driver adaptation to intervention or change. The extent to which this understanding can be described in terms of quantitative relationships will bear on the utility of the research findings. A challenging extension to this research is in the translation of observations of individual user interactions to the scale of an entire roadway or a region. Research components associated with this research theme are described in the remainder of this section.

Driving Simulator and Roadway Data Complementarities

Driving simulators have been used extensively for creating a controlled environment in which to observe driver response to changes in roadway design, operation, or environmental factors. While the natural fidelity of driving simulators has advanced remarkably in recent years, there is still some question as to their ability to capture driver response accurately. The purpose of this research would be to investigate the relation between observations of driver response to changes in roadway design, operation, and environmental factors. These data would be used to improve driver simulation models and validate their predictions.

Models That Capture Micro–Macro Transition: Multiscale Modeling

Traffic safety can be observed and analyzed at a range of physical scales, from the scale of individual drivers (microscopic), to defined road segments and intersections (macroscopic), to large geographic areas (macroscopic). Models applied at each of these scales use different types of predictor variables, partly due to the different phenomena involved, but also because of the type of data observations that are available at each scale. The objective of this research would be to develop systems of models that can transcend these various physical scales and make use of all of the safety information that is available.

Road User Adaptation in Multiple Time Scales: Milliseconds to Months

One of the principal roadblocks for developing sound structural models is the limited amount of knowledge about how road users adapt to interventions and to design and operational changes. Especially important is knowledge about driver adaptation to changes in speed limit and, more generally, driver adaptation to any intervention that is intended to change driver speed. The research would be comprehensive in its investigation of factors that influence road user response and adaptation over time, including how to measure adaptation and how to model it.

SUMMARY

This section provides a summary of the discussion that took place during the workshop and highlights key elements of advanced research for further development of science-based safety evaluation and of more stable, reliable, and transferable highway safety predictive models. The present weakness of safety-related theory is recognized to hamper the further creation of reliable applicable knowledge about cause and effect for road safety management. Workshop participants identified potential research to develop this theory using a structural modeling approach. The research would develop theoretic and empirical relationships that collectively model the driver-vehicle-environment system for the purpose of explaining cause and effect. In this context, theory is meant as any construct that is: (*a*) grounded in (and consistent with) the available crash-based evidence and (*b*) is testable with (and is tested by) crash-based data. This research is envisioned to require the active participation of researchers from multiple discipline areas.

Many participants suggested that surrogate safety measures offer considerable promise and importance in road safety evaluation. What is hindering their application is the sparse research about the link between surrogates and crash-based safety. Whether a certain surrogate predicts crash-based safety is testable. Without such tests, the use of surrogates will remain subject to question. Therefore, research was identified to establish the strength of the relationship between various surrogates and crash-based safety.

To create crash-based information about the effect of interventions and decisions, it is necessary to predict what would have happened if the intervention was not applied, or a different decision made. Since workshop participants raised some question about the accuracy of presently used methods of counterfactual prediction, research was identified to develop better methods for counterfactual prediction.

New and improved methods for the analysis of data will continue to be developed. What is missing is a tool for assessing the extent to which a candidate method succeeds in identifying the cause-and-effect relationship in the data. The development of such a tool is feasible. It would consist of constructing ARD, within which the cause-and-effect structure is known to an independent agency. The data would then be used by researchers who have no knowledge of the cause-and-effect structure. The researchers would apply the method they wish to test and the degree to which they succeed would then be evident. Workshop participants identified potential research to develop and apply the ARD tool was noted.

One of the principal roadblocks for developing sound theories is the weakness of knowledge about how road users adapt to interventions and to design and operation decisions. Especially important is knowledge about speed adaptation and how engineering decisions affect speed choice. Therefore, potential long-term research on adaptation (with priority on speed adaptation) was identified. Due to its fundamental influence on many of the aforementioned research components and theme areas, many participants felt that it would be desirable for research on road user adaptation to be undertaken early on in implementing the advanced safety research activities outlined in this document.

APPENDIX A

Workshop on Future Directions in Highway Crash Data Modeling
Agenda

THURSDAY NOVEMBER, 20, 2008
8:00 a.m.–5:00 p.m.

Welcome
R. Pain, TRB (8:00–8:10 a.m.)

Opening Remarks
J. Toole, FHWA (8:10–8:30 a.m.)

Workshop Overview and Objectives
R. Krammes (8:30–8:45 a.m.)

Highlight workshop objectives and desired product. Review the agenda

Macroscopic Model Building
J. Bonneson (8:45–9:30 a.m.)

Outline the need for and benefits of a sounder theoretical basis underlying macroscopic highway crash data modeling.

Break
(9:30–9:50 a.m.)

Promising Safety Analysis Methods from Allied Fields
R. Krammes, Moderator (9:50–11:30 a.m.; 25 min/speaker)

Cause and Effect in Observational Cross-Section Studies
E. Hauer

Criteria for Causal Inference: Lessons from Epidemiology Applied to Safety Evaluation
R. Elvik

Quantifying Human Performance for Crash Data Modeling
D. Shinar

Lunch
(11:30 a.m.–12:30 p.m.)

Promising Methods in Highway Crash Data Modeling and Analysis
J. Bonneson, Moderator (12:30–1:50 p.m.; 20 min/speaker)

Causal Analysis in Traffic Safety Using Structural Models
G. Davis

The Extreme Value Theory Approach to Safety Estimation
A. Tarko

Modeling Run-Off-Road Crash Frequency and Severity
D. Sicking

Break
(1:50–2:10 p.m.)

FACILITATED DISCUSSION SESSIONS

Part I

Session 1: Guiding Principles (for workshop sessions and product)
Plenary session, facilitated by J. Mason, recorded by J. Bonneson (2:10–2:40 p.m.)

Candidate topic areas for defining the guiding principles:

- Customer;
- Research scope (advanced, applied, synthesis, implementation);
- Relationship to other safety research plans, projects, initiatives;
- Timeframe (short, long, both, ongoing);
- Constraints; and
- Other scope items or assumptions.

Session 2: Reflection, Assessment, and Formulation of Plan
Plenary session, facilitated by J. Mason, recorded by J. Bonneson (2:40–5:00 p.m.)

Discussion objectives:

- Identify and assess promising methods;
- Identify challenges to advancement of methods;
- Formulate a plan for advancing highway safety science; and
- Identify emphasis areas associated with plan and needed advanced research.

Session 3: Breakout groups and topics for Session 3 to be based on the plan emphasis areas and research needs identified in this session.

FRIDAY NOVEMBER 21, 2008
8:00 a.m.–noon

Plan for Second Day
J. Bonneson (8:00–8:15 a.m.)

Organize Breakout Groups for Session 3

FACILITATED DISCUSSION SESSIONS

Part II

Session 3: Breakout Groups to Develop Emphasis Areas and Research Needs
Breakout session, facilitated by designated Group Leader, with designated Recorder (8:15–9:45 a.m.)

Group activities:

- Define vision and goals for assigned topic;
- Identify emphasis area or areas (big-picture, not looking for specific research problem statements); and
- Within each area Identify specific research needs, describe current status, expected results or benefits, relationship to other emphasis areas, timeline, or schedule.

Break
Recorders provide Word or PowerPoint files to J. Bonneson (9:45–10:15 a.m.)

Session 4: Group Review and Organization of Areas and Needs
Plenary session, facilitated by J. Mason, recorded by J. Bonneson (10:15–10:50 a.m.)

Activities:

- Recorder reports by group (2 to 3 min per group);
- Discussion to organize and refine emphasis areas and needs; and
- Discussion of plan for advancement (timeline, management).

Session 5: Group Discussion
Plenary session, facilitated by J. Mason, recorded by J. Bonneson (10:50–11:35 a.m.)

Activities:

- Identify necessary partners (funding agencies, professional organizations, etc.) and
- Next steps.

CLOSURE

R. Krammes (11:35 a.m.–noon)

Summary of Outcomes

Closing Remarks

APPENDIX B

Workshop Participants

James Bonneson
Texas Transportation Institute
j-bonneson@tamu.edu

Kenneth Campbell
TRB
kcampbell@nas.edu

Chou-Lin Chen
National Center for Statistics and Analysis
NHTSA
chou-lin.chen@nhtsa.dot.gov

Forrest Council
Highway Safety Research Center
University of North Carolina, Chapel Hill
f_council@unc.edu

Gary Davis
University of Minnesota
drtrips@umn.edu

Rune Elvik
Institute of Transport Economics
Rune.Elvik@toi.no

Thomas Granda
FHWA
Thomas.Granda@fhwa.dot.gov

David Harkey
University of North Carolina, Chapel Hill
david_harkey@unc.edu

Doug Harwood
Midwest Research Institute
dharwood@mriresearch.org

Ezra Hauer
University of Toronto
ezra.hauer@utoronto.ca

Patricia Hu
Oak Ridge National Laboratory
hups@ornl.gov

John Ivan
University of Connecticut
john.ivan@uconn.edu

Paul Jovanis
Pennsylvania State Univeristy
PPJ2@engr.psu.edu

Ray Krammes
FHWA
Ray.Krammes@fhwa.dot.gov

Dr. John Lee
University of Iowa
jdlee@engineering.uiowa.edu

Neil Lerner
Westat
lernern1@westat.com

John Mason
Auburn University
jmason@engr.psu.edu

Bhagwant Persaud
Ryerson University
bpersaud@ryerson.ca

David Ragland
University of California, Berkeley
davidr@berkeley.edu

David Shinar
Ben Gurion University of the Negev
shinar@bgu.ac.il

Dean Sicking
University of Nebraska
dsicking@unl.edu

Carol Tan
FHWA
Carol.Tan@fhwa.dot.gov

Andrew Tarko
Purdue University
tarko@purdue.edu

Karen Yunk
FHWA
Karen.Yunk@fhwa.dot.gov

Simon Washington
Arizona State University
simon.washington@asu.edu

APPENDIX C

A Cautionary Tale and Promising Directions

Ezra Hauer
University of Toronto

NEED FOR THEORY: A CAUTIONARY TALE

Two motivations have been articulated in this document regarding the need for theoretic explanations of crash events. They can be characterized as:

a. The immediate motivation is the present difficulty of giving causal interpretations to the results of multivariable statistical models when based on available data. The belief is that if one had a theory to guide the selection of structural model equations, then the chance of models to capture cause-and-effect would markedly increase. Only when a model captures cause and effect can one be confident in its predictions about the effect of an intervention or change in design or operation.

b. A more fundamental motivation is the belief that the understanding of accident causes will lead to new and cost-effective interventions.

Failure of Naive Theories for Geometric Design

We should know by now that the making of common sense assumptions about what causes accidents has misled. The theory was that accidents occur on vertical curves (and in other circumstances) when stopping distance (determined by a combination of perception-reaction, friction, and dynamics) is larger than available sight distance. But, on vertical curves sight distance limitation do not seem to be an important cause of crashes. The theory was the wider the lateral separation between oncoming vehicles the fewer accidents will occur. This too is not true beyond a certain lane width. In the formulation of geometric design principles there were theories of sorts (cobbled together from the inheritance of railway design, vehicle dynamics and professional opinion) which did not survive the test of empirical evidence.

Failure of Simple Theories in Ergonomics

There are many tenets, principles and theories which make up the discipline of ergonomics in road safety. One such principle is that making the driving task easier (e.g., enhancing legibility and visibility, improving guidance, reducing workload, etc.) reduces accidents. In road safety this does not seem to pan out. Smiley (*18*) shows instances in which safety was reduced by better path guidance (e.g., through the use of post-mounted delineators, raised pavement markers on sharp curves, and enhanced retroreflectivity) and other instances where safety is improved (e.g., through use of rumble strips). She also shows instances where increased task difficulty improved safety (e.g., when comparing intersections and roundabouts), and where it decreased safety. Thus, the road safety reality is complex and cannot be captured by a few tenets or principles.

Key Difficulty: Adaptation

The key difficulty to developing causal interpretations is adaptive behavior. We have little idea about how to predict the reaction of drivers to what we put in front them. Nor do we know much about how road users behave as a group (e.g., why norms differ from place to place and time to time; to what extent emulation forms norms; how what we do affects norm evolution, etc.). Stated generally, accident occurrence does not depend only on the laws of physics and a few unknown but estimable parameters. It also depends on the complexities of human behavior both as individuals and as groups.

A theory built around physics and probability distributions for parameters will fall short. No theory and no prediction can be on solid ground if it does not entail the complexities of adaptive behavior. Whether the complexities of behavior can be sufficiently understood to become a part of a theory is unclear. The opinion of the committee that put together *Special Report 292* was that, at present, we are quite far from such an understanding even at the immediate level of a driver (*2*).

Should one believe the results of causal models that do not include adaptation? I do not think so. There are too many instances of interventions with consequences that are markedly different from what was anticipated. In most cases the anticipation was frustrated by the complexities of adaptation. It is not as if we have no experience with theories built on opinions, parameters, physics and assumptions. Conclusions based on incomplete theories (the foundations of which are assumptions and opinions) can be misleading and often are.

Common Adaptation: Speed

Driver response in terms of speed is a common and important adaptation with safety consequences. We know that speed changes over time, space, and in response to intervention. However, in spite of generations of research, we know little about what shapes the evolution of speed over time, over space, and how people’s speed choice depends on change. Until we can quantify this response and anticipate its consequences, theories and predictions will be unreliable.

Illustration

Davis (*7*) speaks of a “…unified approach to causal analysis in traffic safety using structural causal models.” The intent is to demonstrate that causal inference is possible when applied to the individual accident and that the results of such inference can lead to the development of accident modification factors (AMFs). His specific question is about the safety effect of reducing the speed limit on a residential street. In making some assumptions about the distribution of vehicle speeds, about what the driver and the pedestrian will do as a function of the distances to the point of collision at the time they see each other, about the distribution of deceleration, it is possible to write equations for the assumed circumstances in which collisions will occur. This, in turn can be plugged into a simulation (using some more assumptions) to obtain a graph of collision probability versus mean speed.

How will people adapt to a reduction from 40 to 25 mph on their street? Will drivers become more alert (because they believe there is a reason why they should slow down) or relax because they believe that since they are driving slowly there is less need to be vigilant? Will parents provide less supervision for children and allow play in the front yard more often because they think the street to be safer? Will children be less careful, etc? I do not know, do you?

Furthermore, the theory does not consider the question of outcome severity. If we know anything about the safety effect of speed it is how it affects severity (and Davis makes use of this knowledge when, later in the same paper, he writes about accident reconstruction). The more severe is an accident, the more likely it is to be reported. Therefore, the number of reported accidents reflects both the probability of accident occurrence and the severity of the outcome. But the proposed theory is confined to modeling only the probability of occurrence; it does not have an element about change in severity. Therefore, its validity cannot be tested by juxtaposition with reported accidents. A theory that cannot be falsified cannot be used for prediction.

Common Element

In the aforementioned illustration, Davis (*4*) reaches conclusions about the relationship of speed and the probability of accidents without reliance on empirical evidence about crashes.

Similarly, Cunto and Saccomanno (*19*), when speaking of simulated vehicle safety performance at signalized intersections, defined a crash potential index (CPI) in terms of the probability that a given vehicle exceeds its maximum available deceleration rate or braking capability for every 0.1 s of simulated time. Unfortunately, the researchers did not see a need to investigate the relationship between CPI and crashes. The common element here seems to be the belief that one can come to conclusions about the safety effect of some action or decision without empirical evidence (direct or indirect) drawn from actual accident occurrence.

A Good Theory Is Good, a Bad One Is Bad

There is promise in developing good theories and danger in applying bad ones. Speculation is not theory; mathematical sophistication does not make a theory. According to the National Academy of Sciences

> In everyday language a theory means a hunch or speculation. Not so in science. In science, the word theory refers to a comprehensive explanation of an important feature of nature that is supported by many facts gathered over time. Theories also allow scientists to make predictions about as yet unobserved phenomena." (*20*).

If road safety management is to be evidence-based, our theories must be rooted in empirical fact and their predictions tested against empirical fact. Hunch and speculations are insufficient. It is in this sense that the risk homeostasis (RH) theory is not a scientific one. First, the soil from which it grew is eclectic, the empirical facts (some of questionable quality) selected because they support the hunch. Empirical facts that did not seem to support the hunch were disregarded. Second, all possible future outcomes can be thought consistent with the RH theory. This makes the theory untestable and its predictions worthless.

For Davis's model to amount to theory one would have to know or show that most collisions between children running into the road and cars in residential areas occur when the parties could not see each other in time (rather than not looking, being distracted, inattentive, tired, drunk etc.). One also would have to know (or show) that the higher the speed limit on residential roads the more numerous are accidents of this type. One would have to ascertain what part of the increase in accidents is attributable to the difference in severity and how much is the increase in frequency. One would have to show that either there is no adaptation to the change in speed limit (in terms of reaction time, attentiveness, play in the front yard, etc.) or know how to predict its extent. To test the theory one would have to make predictions about what change in accidents (net of change due to severity) one should expects for a given change in mean speed and compare this to what has been observed.

In short, theory building requires knowledge of (and testing against) empirical facts which, in the final account, are (reported) crashes.

PROMISING DIRECTIONS

Adaptation: A Missing Piece for Building Theories

Past experience indicates that predictions by road safety theories that did not account for adaptation were often quantitatively (and at times qualitatively) wrong. To succeed in the formulation of road safety theories one must have a grasp on adaptation. Adaptation can take many forms: where, when and by what mode the travel takes place; how fast one drives, how alert one is, etc. Thus research into adaptation should be an important element of future plans. Inasmuch as speed adaptation seems to be a common way in which road users react to change, its examination should be considered first.

Improved Methods of Evaluation

To ground theories in empirical fact and the test them, one needs to do reliable cross-section or before–after comparisons. The central element of all such studies is a prediction of “what would have been” had there been no intervention or difference. Much remains to be done. For cross-section studies, more effort needs to go towards causal structural modeling. For before–after studies, one needs to ascertain which prediction methods work best and devise better methods.

While the empirical Bayes method seems logically attractive and addresses the important regression-to-the-mean issue, there is little empirical evidence to show that it produces better predictions than alternative methods. Nor is there evidence to show that the methods used to develop AMFs and predictive models in the forthcoming HSM produces better predictions than other approaches. In short, fundamental research is needed to advance the art of drawing cause-and-effect conclusions from cross-section and before-after comparisons.

We are fortunate to have cross-section databases like the Highway Safety Information System assembled by FHWA. We should continue to try and exploit these databases. We know by now that cause-and-effect conclusions obtained by methods used until now are unreliable. Better methods may lead to more reliable conclusions. How will we know that conclusions are reliable? The extent to which an approach to modeling succeeds in detecting cause-and-effect relationships is often difficult and at times impossible to determine. However, there may be a way by which to judge the success of a proposed approach. The first step is to create a few ARDs that incorporate plausible and reasonably complex cause-and-effect relationships and noise. Various levels of complexity can be used. The cause-and-effect structure behind the data is kept secret. The next step is to make the data available to researchers who want to try their approach to discovering the causal relationships. The third step is to judge the success of an approach by comparing what the researcher discovered to the assumed causal structure.

Indirect Safety Measurement

The quest for good safety surrogates has great potential. The need for (and promise of) safety surrogates is now greater than ever. The need is greater because the number of potential interventions for which there is no crash data continuously increases (e.g., in-vehicle devices) and the quality of crash data does not seem to be improving (more reliance on self-reporting). The promise is greater because technology and simulation make the collection of surrogates cheaper. However, one must remember that safety is the expected crash frequency by severity. No surrogate can be thought to provide guidance about safety until its relationship with crash data is known and passes the muster of accuracy. In the recent past much effort went into establishing the technological feasibility of cheaply collecting data about surrogates. Little effort went into showing that changes in surrogates and changes in safety go hand-in-hand. It follows that research into the validity of alternative surrogates is now of primary importance.

APPENDIX D

Structural Modeling Approach

GARY DAVIS
University of Minnesota

ISSUES

The methodological approach used in the forthcoming HSM relies heavily on generalized linear models to predict the expected crash frequencies associated with baseline conditions, together with estimates of AMF to capture the effect of modifications from the baseline conditions.

Although plausible, the approach used to develop the HSM models has some limitations that could prevent it from providing a complete solution to the safety prediction problem. One limitation is that it is difficult to capture the interactions between two or more modifications. Another is that, although it is relatively straightforward, given sufficient data, to estimate the AMF associated with the presence or absence of a modification, estimating the AMFs for continuously varying modifications is more difficult. Thus, while estimates have been made of AMFs for installing traffic signals, estimates for changing the length of clearance intervals are more difficult to come by. A third limitation stems from the fact that the baseline models and AMFs are estimated from a sample of roadway sites, and so should be applied only to other sites that are from the same population as those in the sample. Generalizing beyond the model development and validation data sets is done at the user's risk.

Each of the above limitations is, in principle, capable of being overcome, given access to adequate data, and attention to principles of experiment design. A more fundamental limitation is that the current approach is not likely to tell us much about how accidents happen, and this limited understanding will constrain the range of potential modifications that can be identified, and evaluated, in the design process.

One way to understand the importance of this fundamental limitation is to appeal to Paul Thagard's (*15*) thesis about progress in explaining disease. According to Thagard, medical understanding involves three stages. Roughly, it progresses from identifying associations between disease occurrence and other factors, to establishing causal connections, and finally to identifying underlying disease mechanisms. When the mechanism has been identified, the range of possible treatments increases.

While statistical methods are usually central in the first two stages of Thagard's sequence, different methodologies are often needed for the last stage. For example, there is strong evidence supporting the notion that smoking causes lung cancer, and this evidence has been compiled through a series of well-designed, observational studies. However, identifying the mechanism by which smoking causes lung cancer is not a statistical problem. Rather, it is a problem in molecular biology. If the causal process by which cancer cells develop could be specified, the insight could in turn lead to pharmacological interventions to offset disease occurrence.

As another example, in 19th century London, John Snow carried out a classic, observational study to show that the customers of a particular London water company had a higher incidence of cholera than did those of another company (*21*). However, identifying the microorganism that caused cholera was a problem in microbiology, not statistics. Once the microorganism was identified, strategies for preventing its introduction into a water supply, or neutralizing it once it is present, could be devised. If this analogy between traffic accidents and disease holds up, then identifying accident mechanisms should lead to new modifications for preventing accidents, as well as a broader set of tools for evaluating the safety effects of road designs.

At this point, it might be helpful to define some terms. Structural model is taken to mean what Pearl (*22*) calls a causal model, i.e., a set of exogenous variables, a set of endogenous variables, and for each endogenous variable, an equation describing how it changes in response to changes in other variables. If it is possible to specify values for all exogenous variables the model will be deterministic, while if only probability distributions on the exogenous variables are at hand, we have a probabilistic causal model. For a graphical causal model, it is not necessary to specify the structural equations, but only a set of graph nodes, representing the model's variables, and directed arcs and directed arrows, representing causal dependencies.

PROMISING APPROACHES

Structural modeling is not new, and is an active, if rather heterogeneous, area of research. The following list reflects those activities with which the authors are familiar, but is by no means comprehensive. Mayne (*23*) described how a combination of pedestrian and vehicle trajectory models could be used to predict the probability of vehicle–pedestrian collisions, while Baker (*24*) described the use of deterministic models to assess contributions in accident reconstruction. McLean et al. (*25*) reconstructed a number of fatal collisions between vehicles and pedestrians, and then used a variant of Mayne's collision model to predict how different speed management policies might have affected the outcome of each of the reconstructed accidents. This model, in turn, led to an estimate of the safety effect of these policies. Davis (*7*) pointed out how the reasoning underlying this approach could be interpreted as an application of Pearl's causal modeling. More recently, Najm and Koopman (*26*) used a similar approach to evaluate the potential of an Advanced Collision Avoidance System.

Another promising research direction involves combining structural and statistical models. That is, where plausible causal variables can be identified, but where the actual form of the structural equation remains elusive, a statistical model might be used to approximate it. A good example is the work by Kloeden et al. (*27*), where accident reconstruction methods were used to estimate speeds of accident-involved vehicles. These speed estimates were, in turn, used to estimate a logit model relating speed to accident risk. The logit model was used to estimate how the probability of collision might change as speed was reduced. A Bayesian version of this approach was later described in Davis et al. (*28*). Hourdos (*29*) used logit modeling to test for conditional independence in the construction of a detailed graphical causal model for rear-ending accidents on a section of urban freeway.

An interesting finding in Hourdos's work was that a distinctive pattern of traffic conditions, involving speeds and speed differences on adjacent freeway lanes, appeared necessary for the occurrence of rear-end crashes. This finding is interesting because it leads naturally to the issue of surrogate safety measures. A completely satisfying definition of surrogate measures has yet to be given, but arguably a surrogate measure should be an observable traffic condition, that in some sense tracks the occurrence of a given type of accident.

Identification of surrogate measures that have a causal relation to accident occurrence would facilitate the use of traffic simulation models in assessing the safety effects of modifications. The traffic pattern identified by Hourdos could be considered a plausible surrogate measure for rear-end accidents, at least on the section of freeway studied. In this vein, the work by Songchitruksa and Tarko (*30*), using tail probabilities of extreme value distributions to construct surrogate measures, deserves mention.

NEEDED RESEARCH

Before proceeding it may be useful to make a rough distinction between two related, but separate, research activities. The first involves developing tools for supporting practical action. The second involves improving understanding of a phenomenon. Happily, these two activities often complement each other, as practical needs motivate a search for better understanding, which in turn leads to more effective practice. In other situations however these two activities are disconnected, the boundaries between these two activities are poorly understood, and this can result in a tendency to see the methods appropriate for one activity as sufficient for both. This arguably is the case in road safety, and Hauer has given us a forceful account of the risks incurred when prematurely using theoretical reasoning as a decision support tool. The geometric design standard which uses sight distance to constrain vertical curve length is based on a postulated crash mechanism where a driver can only see an obstacle in the roadway when too close to stop without colliding. The unstated assumption is that a substantial fraction of crashes on vertical curves result from this sort of mechanism but, as Hauer (*16*) has noted, this is not necessarily the case, and so there is no demonstrated empirical connection between a supposed safety policy and actual safety. A more subtle distortion can come from the other direction. Statistical modeling of the sort used to support the first edition of the HSM has a clear potential to provide decision-making tools when fundamental understanding of a phenomenon is limited. However, since statistical models assume that the observed phenomena are outcomes of random variables there can be a tendency to project this assumption onto reality, and assume that road crashes are inherently random events (*31*, *32*). The potential danger here is that if road accidents are viewed as objectively random one may conclude that there is no point in proceeding to Thagard's third stage, identification of mechanisms. At least for medical research it is clear that this would have been a mistake.

As noted in Appendix C, mechanism-based methods will require development before they can provide routine decision-support tools. However since near-term practical application is a dominant expectation for road safety research, it may be helpful to consider what is needed to reach this point. Roughly, there are two different, but related, approaches to applying mechanism-based methods to the evaluation of roadway modification (or change). One approach is to start with reconstruction of actual accidents, and then model how the modification would have affected the outcome of each accident. As noted in the preceding section, this approach has seen some limited use (*7*, *25*, *26*), but requires an assumption of how the structural model's background variables will change in response to the intervention. For example, in Davis (*7*) the assumption was that a 25-mph speed limit on a residential street would have no effect on drivers whose chosen speed was less than the new limit, and cause drivers whose preferred speed was greater than the new limit to travel at 25 mph. All other factors affecting an accident were assumed to remain unchanged. What this analysis produces is a ceteris paribus conclusion, e.g., other things being equal what would have been the effect on actually occurring accidents had drivers been constrained to travel no faster than 25 mph. In this case one has actual accidents from which to identify plausible mechanism models, and make estimates of model variables. The primary research need is to determine how road users would respond to the proposed intervention.

Practical difficulties arise with the accident-reconstruction approach because the detailed information needed for reconstruction outstrips what is usually available in a standard accident report. Moreover, the cost of obtaining this information limits its collection to a relatively small number of important cases. It may then be the case that, for the road where design modifications are being considered, no, or very few, investigated accidents are available. One possible solution to this data availability issue would be the routine use of video to record accident events. As Pasanen and Salmivaara (*33*), and more recently, Davis and Swenson (*34*) have shown, it is possible to obtain useful reconstruction data from video recordings of accidents. As noted above though, a more difficult technical challenge is to model how relevant accident variables would change in response to design modifications.

The other approach to applying mechanism-based methods would be to include accident mechanisms as an integral part of traffic simulation models, such as are currently used to evaluate the operational effects of design modifications. Full implementation of this approach will require a reasonably complete identification of types of relevant accident mechanisms, a good understanding of the conditions leading to these mechanisms' activation, and finally, as before, a good understanding of how these conditions will change in response to proposed modifications. The SHRP 2 Safety program may provide at least some of the knowledge needed to implement the simulation-model approach, at least for a limited set of accident types.

APPENDIX E

Artificial Realistic Data
A Research Tool

EZRA HAUER
University of Toronto

In road safety research the true state of affairs is seldom known. This makes it difficult to assess the merit of certain research methods and the validity of some findings. Thus, for example, one does not know the causal dependence of accident frequency or severity on variables. Therefore, one cannot say whether an estimated model equation can serve as a source of AMFs. Similarly, one cannot know the expected proportion of some accident type for a unit. Therefore, it is difficult to know which of the several alternative methods best identifies sites with unusually high proportions. In some circumstances it helps to pretend that the true state of affairs is known and, using these assumed values, to generate an ARD. When the ARD is then subjected to the kind of analysis that would be applied to real data one can ascertain the extent to which what is estimated approximates what was assumed to be true. Three such cases are described in this appendix.

CASE 1: VALIDATE A NOVEL ESTIMATION PROCEDURE OR DETERMINE SAMPLE SIZE

When data come from the laboratory or from randomized trials, estimation and sample size determination are well guided by conventional statistical procedures. In observational studies however, the nature of data and the complexity of circumstances often limit the applicability of conventional statistics. To illustrate, consider the task of using available data to determine the extent by which the retroreflectivity of pavement markings affects the probability of (target) accident occurrence (*35*). The kind of observational data typically available are about accidents (where, when, and of what type), about roads (geometry and traffic) and, in the present context, about markings (marking materials used, when and where applied).

The retroreflectivity of markings declines with the passage of time; if marking retroreflectivity affects accident frequency this might be discernible in the time series of accident counts. The aforementioned complexity of circumstance arises from the fact that pavement remarking is a planned activity and is related to the seasons of the year. Because accident counts also undergo seasonal variation, one has to be able to tell apart the change in the time series of accident counts which is due to the decline in retroreflectivity and that which is due to normal seasonal variation. With this aim in mind, a model representing both time dependencies had to be specified and a procedure for estimating its parameters suggested (*35*). It was not clear whether, with the kind of data that could be obtained, the suggested estimation procedure will 'work'. This is where the ARD became necessary and useful.

The key feature of an ARD is to make plausible assumptions about what is in fact not known. To create this ARD it was assumed that decline in retroreflectivity affects target accident frequency in a specified manner. Using this assumption, a very large ARD (one consisting of a large number of road segments) was then generated and was used to estimate the relationship between retroreflectivity and target accidents. The estimated relationship turned out to replicate well what it was assumed to be when the ARD was generated. This finding was taken to mean that the estimation procedure, as specified, can produce valid results when a large amount of data is available.

Having established that the procedure works, the next task was to determine how many mile-years of observational data are necessary to estimate with sufficient accuracy. This was simple to do. All that was needed was to generate several ARDs differing in mile-years of data and to examine the differences between the estimated and the assumed parameter values. In this manner, the idea of the ARD helped to perform the two essential tasks: (*a*) to prove that the suggested estimation procedure can produce correct results in the real-world setting from which the observational data come, and (*b*) to determine how many mile-years of data are necessary to produce convincing results.

The basic idea is transparent: (*1*) Assume that what is sought is known → (*2*) Using this assumption generates realistic but hypothetical data → (*3*) Use the hypothetical data to estimate what is sought with same methods that you intend to use on real data and → (*4*) Determine whether your results replicate what was assumed to be true and with what accuracy.

To give an impression of what is meant by the phrase "realistic," it is best to describe how, in this case, the hypothetical data were generated. The ARD consists of rows of cells, each row pertaining to one road segment. A road segment is a stretch of road that was remarked as a unit that is, within the same few weeks and by the same material. The essential information about a road segment are the monthly target accident counts for several years such that for each month it is known how much time elapsed since the most recent remarking. This information was generated as follows:

a. It was anticipated that road segments will differ in the number of years for which data are available. To generate this number at random the proportions in Table E-1 were used:

b. To generate at random the road segment length, it was assumed that a typical length remarked under the same contract, by the same material and methods and within a time period of a few weeks is uniformly distributed between 5 to 15 mi.

c. For each road segment and each year, the annual mean number of target accidents had to be generated. Nighttime accidents were thought to be the target. The mean accident frequency was chosen at random from the range 2×10-4 – 3×10-4 per mile and annual average daily traffic (AADT) representing rural two-lane roads. For the first year, the AADT was drawn at random from a uniform distribution in the 5,000 to 15,000 range. It was then assumed that traffic is increasing from year to year by a factor drawn at random from the range 1.00 to 1.03.

d. The annual mean target accident frequency was separated into monthly means using the monthly seasonal factors in Table E-2. These seasonal factors are state-wide averages and do not apply to a specific road segment and year. Accordingly, a random value from a uniform distribution that is 0.2 wide (and centered on the corresponding value in Table E-2) was generated.

TABLE E-1 Proportions of Data-Years

Years of data	1	2	3	4
Proportion of road segments	0.4	0.3	0.2	0.1

TABLE E-2 Seasonal Factors

Months											
1	2	3	4	5	6	7	8	9	10	11	12
1.25	0.97	0.97	0.79	0.84	0.89	0.96	0.98	0.85	0.95	1.15	1.40

Source: E. Hauer, The Effect of Resurfacing on the Safety of Two-Lane Rural Roads in New York State, April 1993.

e. With the results from (*c*) and (*d*), the expected monthly number of accidents as if pavement remarking did not affect accidents frequency was computed.

f. The effect of remarking would be easiest to estimate if it was a year-round activity. Conversely, the effect would be impossible to estimate if all segments were remarked in the same month, for, in that case, the effect of remarking and of the seasonal variation could not be separated. The ARD was generated for two scenarios: (*1*) an unfavorable scenario in which the month when a segment is remarked is uniformly distributed between May 1 and October 1 (5 months) and (*2*) a more favorable scenario in which the month when a segment is remarked is uniformly distributed between March 1 and December 1 (9 months). It was assumed that there are 12 months between consecutive remarkings.

g. The next step was to define how remarking might affect accident frequency. Again two scenarios were considered; one with the assumption that on a freshly remarked segment there are 10% fewer accidents than when the markings are faded, the other assumed that this difference is 5%. Between the fresh and the faded markings the decline in effect of retroreflectivity on accident frequency was assumed to be steep initially and gradual later. These assumptions were used to modify the mean monthly accident frequencies in (*e*) thus accounting for the (assumed) effect of remarking.

h. At this point, we had for each road segment a time series of monthly expected accident frequencies that were shaped by a fairly comprehensive set of variables (number of months of data, segment length, AADT, accident frequency, monthly seasonal factors, the month of remarking, and the assumed remarking effect) such that each variable value was drawn from what was thought to be a realistic distribution.

i. The final step in creating an ARD was to draw Poisson-distributed accident counts from these monthly means.

It would have been difficult to assess the feasibility of unbiased estimation and the accuracy of the estimated parameters by conventional statistical means. The ARD allowed both. It can be done relatively simply and with a degree of verisimilitude that can justify confidence in the results.

CASE 2: COMPARE THE PERFORMANCE OF ALTERNATIVE SCREENING METHODS

At times, the presence of some remediable causal factors becomes manifest at a site as an unusual proportion of accidents of a certain kind. Thus, for example, low pavement friction (or frequent need for braking, or high speeds, etc.) may show up as a higher than average proportion of accidents involving skidding. To identify sites where remedies tailored to reduce specific accident types show promise, the responsible agency usually applies high-proportion screening to all sites.

There are several alternative methods for high-proportion screening. When considering which method to use, the better screening method is that which has near the top of its ranked list the larger number of sites where the proportion is truly high. The difficulty is that, in reality, the true proportion for a site is never known, and therefore, a clear proof of which screening method is best is difficult to provide. Here again, the ARD artifice becomes useful. To create this ARD one begins by assuming realistic expected frequencies of 'target' (μ_t) and of 'other' (μ_o) accidents at many sites. The ratio $\mu_t/(\mu_t + \mu_o)$ is the true proportion of target accidents for a site. Using the two expected accident frequencies μ_t and μ_o accident counts are generated at random from the corresponding Poisson distributions. The alternative screening methods are then applied to this ARD. There are as many ranked lists as there are methods of screening. Since now the true proportion for each site is known, it is easy to determine the top of which ranked lists contains a richer distribution of true proportions.

CASE 3: ASSESS METHODS OF MULTIVARIATE MODELING

There are many ways to fit a multivariate statistical model to data. The ways differ in the data used, in assumptions made about the error structure, in decisions about which variables to include in the model equation and which to exclude, in the choice of the mathematical function to fit to the data, in the optimization criterion used, etc. Most fitted models do a good job of estimating the expected accident frequency. However, most seem to do a questionable job of predicting the safety consequences of manipulations.

For an equation to predict the consequences of manipulations, the expression on its right-hand side must stand in a cause-effect relationship with the response variable on the left-hand side. The difficulty resides in the fact that multivariate statistical models represent the associations in the data (covariation, correlation) and associations exist for many reasons, only one of them being the relationship between cause and effect.

Because of the widespread tendency to use the results of multivariate statistical modeling as if they did predict the effect of manipulations, it is important to ask which of the many ways of multivariate modeling comes closer to representing cause and effect and whether the best modeling approach comes close enough. As the true causal linkages are not known, this all-important question is difficult to answer. Once again the use of an ARD may help.

The penultimate step in ARD creation is to construct an expected accident frequency for each unit. This expected accident frequency is the point of departure for the final step: the creation of accident counts. How the expected frequency needs to be constructed depends on the questions which the ARD is supposed to answer. In the present case the main question is about the identification of the causal structure that gave rise to the data. Accordingly, the expected accident frequencies in the ARD have to be built to realistically represent the principal obstacles that hide the causal structures from view. What obstructs our ability to detect causal structures in observational data is fairly well understood. The major difficulties are:

a. The use of aggregate variables and averages. For example, modeling of total accidents instead of separate accident types; representing traffic by AADT (which is the average for a representative day of a year) rather than, say, by hourly volumes; representing obstacles by the roadside hazard rating (which is an average impression for the entire segment) instead of individual obstacles; or the use of driveway density (instead of individual driveways by type).

b. Errors in variable (e.g., the AADT estimates are factored up from a few days of counts conducted once every few years and interpolated for years in which no counting was done).

c. The common absence of important causal variables [e.g., precipitation; speed distribution; pavement friction; driver age, gender, income, and blood alcohol content (BAC); vehicle mass, vintage, and occupancy; quality of emergency medical services, distance from hospital, etc.]

d. The presence of complex dependencies amongst variables (e.g., AADT affects the choice of design speed which in turn influences and various geometric design elements and thereby affects road user behavior and speed choice. In addition, AADT both reflects and influences adjacent land use and access patterns).

e. The use of simple mathematical functions (usually without interactions) to represent the complexities and intricacies of causal webs.

An ARD (the purpose of which is to answer the question about how well a modeling approach can identify specified cause-effect relationship) must address points (*a*) to (*e*). That is:

- Ad (*a*). The ARD should contain the disaggregate values to which the causal relationships can be attributed and from which the aggregates and averages used for modeling can be computed. Thus, for example, the probability of a single-vehicle accident to occur may more clearly depend on the average hourly flow by time of day than on the average daily flow. Therefore, the ARD should perhaps contain the 24-h profile of flows for each day of the year.
- Ad (*b*). When the ARD is used to create the data for modeling, appropriate uncertainties and errors should be created. Thus, for example, to create the AADT for use in modeling, one might take a sample of few days from the detailed traffic profile, introduce some counting errors, and the usual expansion factors to create an AADT estimate for the years in which a count is conducted and an interpolated value for the years without counts.
- Ad (*c*). The ARD should be built of most variables that influence accident occurrence and severity, not only of those for which data can be obtained. All these variables will serve to generate the expected accident frequencies and severities. However, only some of this set will serve to generate the explanatory variable data used for modeling. This feature of the ARD will enable the researcher to examine the importance of missing variables. Suppose, for example, that higher-than-average BAC content is found to be associated roads that have lower-than-average geometric standards and where hospitals and ambulance stations are rare. The ARD should be built with BAC distribution and hospital access time as causal variables. The absence of these from the modeling data will cause a bias in the influence of geometric variables.
- Ad (*d*). There are many causal and correlational associations amongst the explanatory variables. Thus, for example, the speed of travel depends on the geometry of the road and on its traffic. This makes it difficult to separate by statistical means the causal effects speed geometry and traffic. Similarly, as noted earlier, AADT is a causal influence not only of accident occurrence but also on many variables (e.g., the presence of shopping plazas, the standard of winter maintenance, or the intensity of police enforcement). To be realistic, the ARD must capture the complex relationships between the explanatory variables.

- Ad (*e*). In the ARD, one should represent a plausible, realistic and detailed causal structure. The many complexities and intricacies of the causal web will become evident when detailing the ARD.

The construction of an ARD of this kind requires much thought and is likely to be a significant undertaking. The reward is that important questions that were heretofore without answer may become answerable.

The main question is about the modeling tools with which casual questions of interest can be sufficiently reliably answered. However, there are many other questions which the ARD can answer: What variables can be aggregated and to what extent? Which variables are essential and which of secondary importance? What sample size is required?

References

1. Hauer, E. *Observational Before–After Studies in Road Safety*. Elsevier Science, Inc., Tarryton, N.Y., 1997.
2. *Special Report 292: Safety Research on Highway Infrastructure and Operations: Improving Priorities, Coordination, and Quality.* Transportation Research Board of the National Academies, Washington, D.C., 2008.
3. Evans, L. The Dramatic Failure of U.S. Traffic Safety Policy: Engineering is Important, Public Policy is Crucial. *TR News*, No. 242. Transportation Research Board, National Research Council, Washington, D.C., January–February 2006, pp. 28–31.
4. Davis, G. Possible Aggregation Biases in Road Safety Research and a Mechanism Approach to Accident Modeling. *Accident Analysis and Prevention*, Vol. 36, 2004, pp. 1119–1127.
5. Hauer, E. Fishing for Safety Information in Murky Waters. *Journal of Transportation Engineering*, Vol. 131, No. 5, 2005, pp. 340–344.
6. Hauer, E. Cause and Effect in Observational Cross-Section Studies on Road Safety. Unpublished paper. 2005. Available at http://www.roadsafetyresearch.com/March 2009.
7. Davis, G. Towards a Unified Approach to Causal Analysis in Traffic Safety Using Structural Causal Models. *Transportation and Traffic Theory in the 21st Century* (M. Taylor, ed.), Elsevier Science, Inc., Tarryton, N.Y., 2002, pp. 247–266.
8. Anastasopoulos, P., and F. Mannering. A Note on Modeling Vehicle Accident Frequencies with Random-Parameters Count Models. *Accident Analysis and Prevention*, Vol. 41, 2009, pp. 153–159.
9. Shugan, S. Editorial: Errors in the Variables, Unobserved Heterogeneity, and Other Ways of Hiding Statistical Error. *Marketing Science*, Vol. 25, No. 3, 2006, pp. 203–216.
10. Elvik, R. Developing Accident Modification Functions: Exploratory Study. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2103*, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 18–24.
11. *Highway Safety Manual*. AASHTO, Washington, D.C., 2010.
12. Shinar, D. Speed and Crashes: A Controversial Topic and an Elusive Relationship. *Special Report 254: Managing Speed: Review of Current Practice for Setting and Enforcing Speed Limits*, TRB, National Research Council, Washington, D.C., 1998, pp. 221–276.
13. McCarthy, P. Effect of Speed Limits on Speed Distributions and Highway Safety: A Survey of the Literature. *Special Report 254: Managing Speed: Review of Current Practice for Setting and Enforcing Speed Limits*, TRB, National Research Council, Washington, D.C., 1998, pp. 277–358.
14. Elvik, R. To What Extent can Theory Account for the Findings of Road Safety Evaluation Studies? *Accident Analysis and Prevention*, Vol. 36, 2004, pp. 841–849.
15. Thagard, P. Explaining Disease: Correlations, Causes, and Mechanisms. *Minds and Machines*, Vol. 8, 1998, pp. 61–78.
16. Hauer, E. The Engineering of Safety and the Safety of Engineering. *Challenging the Old Order: Towards New Directions in Traffic Safety Theory* (J. Rothe, ed.), Transaction Publishers, London, 1990, pp. 29–71.
17. Davis, G. Accident Reduction Factors and Causal Inference in Traffic Safety Studies: A Review. *Accident Analysis and Prevention*, Vol. 32, 2000, pp. 95–109.
18. Smiley, A. Driver Behavior: A Moving Target. *TR News,* No. 254, Transportation Research Board, National Research Council, Washington, D.C., 2008, pp. 19–24.
19. Cunto, F., and F. Saccomanno. Calibration and Validation of Simulated Vehicle Safety Performance at Signalized Intersections. *Accident Analysis and Prevention*, Vol. 40, 2008, pp. 1171–1179.
20. Science, Evolution, and Creationism. Brochure adapted from *Science and Creationism: A View from the National Academy of Sciences, Second Edition*, Institute of Medicine, National Academy of Sciences, Washington, D.C., 2008.

21. Freedman, D. From Association to Causation: Some Remarks on the History of Statistics. *Technical Report 521*, University of California–Berkeley, 2002.
22. Pearl, J. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.
23. Mayne, A. The Problem of the Careless Pedestrian. *Proc., Second International Symposium on the Theory of Road Traffic* (J. Almond, ed.), OECD, 1965, pp. 279–285.
24. Baker, J. *Traffic Accident Investigation Manual*, Northwestern University Traffic Institute, 1975.
25. McLean, J., R. Anderson, M. Farmer, B. Lee, and C. Brooks. Vehicle Travel Speeds and the Incidence of Fatal Pedestrian Crashes, Road Accident Research Unit, University of Adelaide, 1994.
26. Najm, W., and J. Koopman. Performance of a Rear-End Crash Avoidance System in a Field Operational Test. SAE Technical Paper 2006-01-0573, 2006.
27. Kloeden, C., J. McLean, and G. Glonek. *Reanalysis of Traveling Speed and the Risk of Crash Involvement in Adelaide South Australia*, Road Accident Research Unit, University of Adelaide, 2002.
28. Davis, G., J.-P. Pei, and S. Davuluri. Speed as a Risk Factor in Serious Run-Off-Road Crashes: Bayesian Case-Control Analysis with Case Speed Uncertainty. *Journal of Transportation and Statistics*, Vol. 9, 2006, pp. 17–28.
29. Hourdos, J. Crash-Prone Traffic Flow Dynamics: Identification and Real-Time Detection. PhD. thesis, University of Minnesota, 2005.
30. Songchitruksa, P., and A. Tarko. The Extreme Value Theory Approach to Safety Estimation. *Accident Analysis and Prevention*, Vol. 38, 2006, pp. 811–822.
31. Hauer, E. Traffic Conflicts and Exposure. *Accident Analysis and Prevention*, Vol.14, 1982, pp. 359–364.
32. Lord, D., S. P. Washington, and J. N. Ivan. Statistical Challenges with Modeling Motor Vehicle Crashes: Understanding the Implication of Alternative Approaches. Presented at 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2004.
33. Pasanen, E., and H. Salmivaara. Driving Speed and Pedestrian Safety in the City of Helsinki. *Traffic Engineering and Control*, June 1993.
34. Davis, G., and T. Swenson. Collective Responsibility for Freeway Rear-Ending Accidents? An Application of Probabilistic Causal Models. *Accident Analysis and Prevention,* Vol. 38, 2006, pp. 728–736.
35. Masliah, M., G. B.Bahar, and E. Hauer. Application of Innovative Time Series Methodology to Relationship Between Retroreflectivity of Pavement Markings and Crashes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019*, Transportation Research Board, Washington, D.C., 2007, pp. 119–126.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. (Dan) Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. (Dan) Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

TRB
TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001