UNIVERSITY OF MINNESOTA

CENTER FOR TRANSPORTATION STUDIES

INTELLIGENT TRANSPORTATION SYSTEMS INSTITUTE

A Case Control Study of Speed and Crash Risk

Technical Report 1

Aggregation Biases in Road Safety Research and a Mechanism Approach to Accident Modeling

Final Report

Prepared by:
Gary A. Davis

Department of Civil Engineering
University of Minnesota

CTS 06-01A

HUMAN CENTERED TECHNOLOGY TO ENHANCE SAFETY AND TECHNOLOGY

Technical Report Documentation Page

1. Report No. CTS 06-01A	2.	3. Recipients Accession No.
4. Title and Subtitle A Case Control Study of Speed and Crash Risk Technical Report 1 Aggregation Biases in Road Safety Research and a Mechanism Approach to Accident Modeling		5. Report Date April 2006 6.
7. Author(s) Gary A. Davis		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil Engineering University of Minnesota 500 Pillsbury Dr SE Minneapolis, MN 55455		10. Project/Task/Work Unit No. CTS Project Number 2001032 11. Contract (C) or Grant (G) No.
12. Sponsoring Organization Name and Address Intelligent Transportation Systems Institute Center for Transportation Studies University of Minnesota 511 Washington Avenue SE, Suite 200 Minneapolis, MN 55455		13. Type of Report and Period Covered Final Report December 2000-February 2002 14. Sponsoring Agency Code
15. Supplementary Notes http://www.cts.umn.edu/pdf/CTS-06-01A.pdf		

16. Abstract (Limit: 200 words)

In accident reconstruction, individual road accidents are treated as essentially deterministic events, although incomplete information can leave one uncertain about how exactly an accident happened. In statistical studies, on the other hand, accidents are treated as individually random, although the parameters governing their probability distributions may be modeled deterministically. Selection of one or the other of these approaches affects how data are interpreted, and here a simple deterministic model of a vehicle/pedestrian encounter is used to illustrate how naively applying statistical methods to aggregated data could lead to an ecological fallacy and to Simpson's paradox. It is suggested that these problems occur because the statistical regularities observed in accident data have no independent status but are simply the results of aggregating particular types and frequencies of mechanisms.

17. Document Analysis/Descriptors		18.Availability Statement	
Speed limits, accident analysis, statistical model, mechanism model, human factors in accidents, traffic safety		No restrictions. Document available from: National Technical Information Services, Springfield, Virginia 22161	
19. Security Class (this report) Classified	20. Security Class (this page) Classified	21. No. of Pages 24	22. Price

A Case Control Study of Speed and Crash Risk
Technical Report 1
Aggregation Biases in Road Safety Research and a Mechanism Approach to Accident Modeling

Final Report

Prepared by:
Gary A. Davis

Department of Civil Engineering
University of Minnesota

April 2006

Intelligent Transportation Systems Institute
Center for Transportation Studies
University of Minnesota

CTS 06-01A

Acknowledgements

The author wishes to acknowledge those who made this research possible. Funding for the study was provided by the Intelligent Transportation Systems (ITS) Institute at the University of Minnesota's Center for Transportation Studies (CTS). The ITS Institute is a program funded by the United States Department of Transportation (USDOT) and administered by the USDOT's Research and Innovative Technologies Administration (RITA).

A Case Control Study of Speed and Crash Risk

Technical Report 1

Aggregation Biases in Road Safety Research and a Mechanism Approach to Accident Modeling

Gary A. Davis

University of Minnesota

Introduction

More than a dozen years ago, Ezra Hauer pointed out that while a stated aim of road design practice was to provide safe travel, the empirically-based knowledge that would allow an engineer to predict the safety consequences of design decisions was, in many cases, nonexistent (Hauer 1990). The volume of literature published since then suggests that Hauer's challenge has not been ignored. Since 1991 the annual meetings of the Transportation Research Board have, each year, included between 50 and 100 papers and presentations related to "Safety and Human Performance," while a TRIS search using the keywords "highway safety" yielded over 1900 items published after 1990. In the United States, efforts to develop a Highway Safety Manual similar in spirit to the Highway Capacity Manual, and the Federal Highway Administration's Interactive Highway Safety Design Model (IHSDM), were begun during the 1990s. But it is less clear that this quantity of effort has led to a proportionate advance in knowledge. For example, in 1985 the U.S. Congress relaxed the National Maximum Speed Limit (NMSL) on rural freeways, and in 1995 repealed the NMSL entirely. However, one can find papers reporting increases, no change, and even decreases in accidents attributable to increases in speed limits, and there is at present no consensus on the effects of the speed limit repeal. Two reviews of literature on speed and safety, commissioned by the Transportation Research Board (Shinar 1998; McCarthy 1998), respectively cited 73 and 65 sources before concluding that although evidence tended to support the notion that accident risk increased with speed, more study was needed to identify those conditions where changes in speed limit can affect accidents, or to predict the sizes of these effects. As another example, over the past decade a number of studies have used statistical models to correlate accident experience with variations in traffic and road conditions. Recently, though, studies assessing the transferability of such models have found that the set of variables identified as significant accident predictors can differ for data collected in the same geographic region but at different times, as well as for data collected in different regions (Persaud et al. 2002; Lyon et al. 2003). So despite their frequent application, the ability of such models to reliably identify important accident predictors is open to question.

The physical and social conditions that produce road accidents are obviously complex, so it is perhaps not surprising that untangling the causal contributions of these conditions has proved difficult. This is even more so because the study designs available to accident researchers are almost always observational, rather than experimental. I would like to suggest, though, that at least some of the present uncertainty may be due as much to characteristics of a commonly-used research approach as to the inherent complexity of accident phenomena. The main feature of this approach is the use of statistical models to estimate what Turner (1997) calls "net effects." The expectation is that by using a statistical model to correct for confounding influences, the causal effect of some factor of interest can be isolated. An advantage of this approach is that one need only identify possible influences on accident phenomena, and can then use data and statistical analysis to convert this to a quantitative understanding. However the extent to which, or even if, this approach can succeed is a topic of active debate (e.g. McKim and Turner 1997). In what follows, I will first contrast a view on the nature of road accidents sometimes used to support the use of statistical modeling, with an alternative view which appears to underpin the investigation of individual accidents. I will then illustrate how these views can lead to nontrivial differences in the interpretation of accident data. Finally, in the Conclusion section I will attempt to relate this discussion to some broader issues.

Foundations for Inferences about Accidents

Many of us manage to apply statistical methods in our research without worrying overmuch about the foundations of statistics. Still, a statistical analysis is essentially a logical argument, where assumptions about the processes generating data are combined with observation statements in order to derive statements about quantities not directly observed. Inconsistencies between what has been assumed and how the data have actually been produced could then invalidate any conclusions drawn from the analysis. In this section I would like to make a rough comparison between what many would take to be the underlying assumptions of the statistical approach to accident analysis (e.g. Lord et al. 2004), and an alternative set of assumptions drawn from work on foundations for causal inference. A basic notion introduced in many statistics texts is that of a random experiment, which is an observation activity that can be repeated under identical conditions, where the set of possible outcomes is known, but where the outcome of any particular repetition is not known in advance (e.g. Rohatgi 1976, p. 20). Hacking (1964 1965) has given a treatment of the foundations for statistical inference that explicates this idea, and begins by noting that our world contains devices or situations which appear to produce, under repeated operation, stable relative frequencies of outcomes, even though the actual outcome of a particular operation is unpredictable. That is, there exist devices or situations which correspond to the formal notion of a random experiment. Hacking calls this tendency to produce stable relative frequencies from unpredictable individual outcomes, chance, and the corresponding situation or device is called a chance set-up. Probability theory then provides a logic for reasoning about chance. Hacking takes it as given that chance set-ups are part of the furniture of our world, and that the chance of an outcome is an objective property of a particular set-up. The classic example of a chance set-up is tossing a coin, but Hacking also suggests that traffic accident frequencies can be considered as resulting from chance set-ups. A goal of statistical inference is then to use observations on a finite number of trials of a set-up to estimate the chances of particular outcomes, and to determine how these chances vary as features of the set-up are changed.

As Hauer (1982) has noted, one can readily apply Hacking's treatment to the study of road accidents by assuming, in the simplest case, that a section of road or an intersection can be modeled as a chance set-up. Individual vehicle traversals provide the trials to which the set-up is subjected, each trial is assumed to have a chance of resulting in an accident, and this chance may vary with roadway, traffic, or environmental conditions. This leads to two related practical advantages. First, by observing how accident frequencies vary across different sites, or in response to modifications at given sites, the tools of statistical analysis can be used to study empirically how decisions concerning road design and operation affect the chances of accidents. Second, just as it is possible to use a sequence of coin tosses to estimate the chance of the coin coming up heads without knowing the details of any particular toss, so it should be possible to use accident statistics to study the chances of accidents without knowing the details of any particular accident. This in turn supports using research approaches for which the computerized accident records maintained by public agencies are a primary data source. An important point to note is that although one might conclude from a statistical study that an intervention has lowered the chance of accidents at a site, one cannot say that the intervention prevented (or caused) any particular accident. The intervention affects the chance, but then the individual accidents either do or do not happen (Glennon 1997).

If statistical methods were necessary to identify the causes of accidents, then those types of accidents which happen too infrequently to permit stable estimates of chance would be a *terra incognita.* In particular, since accidents involving large passenger aircraft or passenger trains are rare, it would not be possible to discover the causes of these accidents. But this contradicts the practice of agencies such as the National Transportation Safety Board, which carries out detailed investigations of individual accidents in order to determine their probable causes and to make recommendations for preventing similar accidents in the future (Kapustin 2001). When we turn to road accidents, it is sometimes the case that an individual accident is investigated in detail, and then it is sometimes possible to model the accident as the outcome of a deterministic mechanism, in the sense that a set of structural equations together with an assignment of values to variables describing particular driver, vehicle, and road characteristics, can be used to retrodict the evidence from that accident. What makes such reconstruction possible is that transportation accidents usually involve human-made systems, which have been designed and built to behave (approximately) like deterministic mechanisms. This means that rather than starting from a state of ignorance, the accident investigator can often draw on background knowledge concerning how a particular system was designed to behave, or how similar systems have behaved in the past. The data needed to carry out such investigations and reconstructions are, however, much more extensive than what are available in a typical computerized accident record.

To put the issue baldly, the statistical approach to investigating accidents appears to assume that road accidents are individually unpredictable, chancy phenomena, although aggregates of accidents can show predictable statistical regularities. The reconstruction approach on the other hand assumes that accidents are the result of deterministic processes, although the provenance of any particular accident may be uncertain because of limits on the information available. We appear then to have contradictory views on the basic nature of road accidents, as contradictory as the notions of chance and determinism. I hope that I am not alone in finding this situation unacceptable. If this appearance of conflict is due to the state of our knowledge rather than to some fundamental duality in the nature of accidents, it should be possible to develop a foundational treatment that includes both statistical studies and accident reconstruction. Fortunately (or unfortunately) this issue is not peculiar to accident research, and has arisen in other fields where a primary aim is to make causal inferences about complex, partially understood systems. Like many important ideas, variants of this approach have been given by different authors. In statistics, there is Holland's (1986) explication of the Rubin's potential response model (Rubin 1974 1978). For more philosophical treatments, there is Giere's (1980) account of the meaning of causal hypotheses in statistics, and more recently Glennon's (1996 1997) "analysis of causation based on a theory of mechanisms" (1996, p. 49). Schaffner (1993) has reviewed much of the earlier work on this issue while attempting to identify commonalities in the notions of causation used in laboratory sciences, such as physiology, and those used in epidemiology. Finally, an impressive attempt to unify these ideas, and to convert them from philosophical curiosities to practical tools, has been recently given by Pearl (2000). This simplifies our task. Rather than having to develop an alternative foundational treatment from scratch, we can first try to apply these ideas to road accidents.

In its simplest form, this alternative treatment begins with the idea of a population of individual deterministic mechanisms, rather than with the idea of repeated trials of a chance set-up. Each of these mechanisms has a well-defined response to a causal intervention, but because of individual differences the mechanisms do not necessarily respond identically. Subjecting all members of this population to the same level of a causal intervention would in principle produce one distribution of responses, while subjecting all members to a different level would produce a possibly different distribution. One can then use these hypothetical distributions to define what is meant by an aggregate causal effect, even if the details of how the intervention operates on an individual mechanism remain vague. A basic problem of causal inference then is to reliably estimate these effects from restricted samples. As with Hacking's treatment it may not be possible to predict how any particular individual mechanism will respond, but this uncertainty is now attributed to ignorance of the details of the individual case rather than to the workings of objective chance.

This potential response model provides a foundation for statistical inference especially applicable to problems involving causal inferences from observational data. A criticism leveled against it however concerns the nature of the potential responses, holding that the individual responses to interventions which do not actually materialize are not well-defined, and so vulnerable to subjective interpretation (Dawid 2000). Pearl (2000) points out however that this difficulty arises mainly because the potential responses in Rubin's model are treated as black boxes. If instead one has a plausible structural equation model of how the mechanisms work, then the potential response of a mechanism, given specified background conditions, is well-defined and can be found by solving the system of structural equations. Using a probability distribution to summarize how the input values are distributed across the population then induces a probability distribution over the mechanisms' responses, but this does not mean that the responses are due to objective chance. Given a full specification of input values the responses become deterministic. Pearl's approach also implies that the statistical and the reconstruction approaches to studying road accidents are more like different points on a continuum, distinguished by the level of detail used to describe each accident, and the number of accidents included in the analysis. More importantly, while under the chance set-up view a statistical estimate of an intervention effect could be interpreted as measuring how an objective quantity, the chance of an accident, changed in response to the intervention, under this individual mechanism approach statistical assessments of causal effects have no independent existence. Rather, they are simply summaries of how the intervention affected a set of individual accident events. In principle then it is possible, depending on how the individual events are selected and aggregated, to obtain statistical measures that are at odds with what is happening at the individual level. In what follows I will illustrate how this could occur with regard to a relation between vehicle speed and pedestrian accident risk, and with regard to the effect of a speed limit reduction.

Potential Ecological Bias in Statistical Accident Models

One way in which studies of speed and accident risk tend to differ is in what they take as the basic units of analysis. For example, in Kloeden et al (1997) estimates of the speeds of individual vehicles involved in accidents were compared to the speeds of comparable control vehicles, while Garber and Gadiraju (1988) used highway segments as basic units of analysis, and Lave (1985) used the States of the Union. In what follows I will use a model of an accident mechanism to illustrate how a clear relationship between speed and individual accident risk can be obscured in an aggregated cross-sectional study. To do this, I will draw on a study described in more detail in Davis et al (2002). The objectives of this study were to develop a method for prioritizing residential streets with respect to their appropriateness for neighborhood traffic control (traffic calming), and to predict the safety effects of a hypothetical 25 mph (40 km/h) speed limit. Because pedestrian accidents happen infrequently on such streets, standard statistical methods were not applicable. Hypothetically, one could instead subject each street of interest to a standardized test, in which a set of standardized pedestrians each enter the street, and the number of these being struck by cars is recorded. This obviously is also not feasible, so the chosen approach was to simulate the outcome of such a test, in which heedless pedestrians run into the street and then stop in the traveled way. This simulation model can be used to produce hypothetical data from three levels of aggregation, the individual vehicle/pedestrian encounter, the population of encounters at a given site, and the population of sites. The individual encounters were characterized using the pedestrian's speed (v_2) and initial distance (x_2), the vehicle's initial speed (v_1) and distance (x_1), and the driver's reaction time (t_p) and braking rate (a). The outcome of the encounter was simulated using simple kinematics, such that if the vehicle could pass the collision point before the pedestrian arrives, $(x_2/v_2)>(x_1/v_1)$, or if the driver could stop before reaching the collision point, $x_1>v_1t_p+v_1^2/2a$, the collision was avoided (y=0), otherwise it occurred (y=1). By placing probability distributions on these variables it was then a simple task to compute the collision probability P[y=1] using Monte Carlo simulation. As part of the study, portable traffic counters were used to collect vehicle speed and headway data from 25 residential streets, and average building setbacks for each of these streets were measured from aerial photographs. These data, together with reference distributions for v_2, t_p, and a were then used to compute the collision probabilities during peak-hour conditions on each street. Table 1 summarizes the characteristics of the probability distributions used, while Figure 1 displays the resulting collision probabilities plotted against average peak-hour speed, and peak-hour traffic volume.

The research team then looked into trying to simplify this approach by using a statistical model to relate the collision probabilities to aggregate summaries of the traffic variables, and so avoid the need to carry out Monte Carlo simulations. A logistic regression of the collision probabilities versus traffic volume and mean speed produced the following estimated equation

$$L(P[y=1]) = -4.82 + (.018)\,(\text{volume}) + (.0053)(\text{mean-speed}) \qquad (1)$$

where L(p)=log(p/(1-p)) denotes the logistic transformation. The standard errors for the volume and mean-speed coefficients were .003 and .015 respectively. It is straightforward to verify that while collision probability is clearly associated with traffic volume, the association with mean speed is noticeably weaker, to the extent that a model using only traffic volume as its independent variable would predict nearly as well as one which also included mean speed. This in turn appears to suggest that, at least for the range of speeds that appear to be typical for residential streets, speed need not be considered in assessing pedestrian risk.

This is puzzling, because at the level of the individual pedestrian/vehicle encounter, whether or not a simulated collision occurs is a deterministic function of the simulation model variables x_1,v_1,x_2,v_2,t_p and a. For most reasonable values of x_1,x_2,v_2, t_p, and a, y is a unit step function of v_1, with the jump point being determined by the other five variables. Furthermore, at an intermediate level of aggregation, that of the individual site, a relationship between mean vehicle speed and collision probability can be determined by repeated simulations, varying the mean vehicle speed but keeping all other distribution parameters fixed. Figure 2 shows the resulting relationships for two sites, numbers 22 and 27b, where it can be seen that collision probability increases roughly linearly as mean speed increases, although the slopes of these lines differ for the two sites. These results are characteristic of what was observed for the other sites.

So, if at the level of the individual collision, and at the level of the individual site, there is an increasing relationship between traffic speed and pedestrian accident risk, why isn't this relationship present when the site data are aggregated? The answer becomes clear once one understands what happens at the individual sites. As Figure 2 shows, collision risk increases approximately linearly with mean speed at each site, and it turns out that the slope of this line is roughly proportional to the traffic volume. This particular set of locations showed a wide range in traffic volumes (the site with lowest volume and that with the highest volume differed by a factor of 10), but a fairly narrow range in average speeds (these differed by at most a factor of 2). This produced the pattern shown in Figure 1, where the interaction between volume and mean speed, together with traffic properties of this particular sample of sites, combined to make the collision probability appear to vary randomly at each speed level. Knowing this, one could suggest that a more systematic sampling procedure, which first stratified by traffic volume and then at each volume level selected sites showing a range of mean speeds, would have been more likely to identify the relationship between mean speed and risk shown in Figure 2.

Freedman (1999a) has defined an ecological fallacy as concluding that a relationship observed in aggregated data necessarily holds at disaggregated levels, and for this example, concluding from the aggregate analysis that pedestrian risk is essentially independent of average speed would be an ecological fallacy. Yet since a more appropriately structured sample could have identified the relationship between speed and risk, aggregated analyses do not necessarily produce misleading results. The problem is that without prior knowledge concerning the underlying mechanism generating the aggregated data it may be difficult, or even impossible, to correctly interpret the aggregated results. Certainly if the objective is to identify those structural features which cause traffic accidents, the common practice of fitting regression-type models to convenience samples of accident data should be considered suspect. This point has been made elsewhere (Turner 1997; Freedman 1997).

Simpson's Paradox and Before/After Studies

As noted in the Introduction, another feature of the research on speed and accident risk is the lack of consensus shown by studies investigating how changing a speed limit affects accident occurrence. The classic Simpson's paradox (Simpson 1951) arises in the interpretation of contingency tables, when an association between variables observed in sub-populations is attenuated or even reversed when the sub-populations are aggregated. In this section, the pedestrian collision model described earlier will be used to illustrate how a similar effect might occur when one attempts to estimate the effect of a lowered speed limit on pedestrian safety. In the sample of residential streets described earlier, all the streets had nominal limits of 30 mph (48 km/h) and one of the objectives of the study was to simulate the potential effect of reducing this limit to 25 mph (40 km/h). To see how this can be done, some background on characterizing causal effects will be helpful.

A commonly-used measure of the effect of a safety intervention is its reduction factor (RF), which in the simplest case can be defined operationally via an expression such as

$$RF = (P[y=1|z=0]-P[y=1|z=1])/P[y=1|z=0] \qquad (2)$$

Here z=0 stands for absence of the countermeasure of interest, while z=1 stands for its presence. Given accident counts and exposure measures at locations and/or times with and without the countermeasure, one can compute estimates of the conditional probabilities appearing in equation (2), and in turn estimate the RF. Hauer (1980) has pointed out, though, that equations such as (2) cannot be taken as defining what is meant by the RF, since application of the estimator to data subject to selection biases can give a biased estimate of the RF. Rather, the reduction factor is better understood as an underlying property of the countermeasure, which equation (2) may, or may not, estimate reliably. In a more general context, Pearl (2000) calls this underlying property the probability of necessity (PN), which can be interpreted as the fraction of accidents which occurred when the countermeasure was absent that would have been prevented had the countermeasure been present. Pearl also gives a theorem showing that when certain conditions (called monotonicity and exogeneity) are satisfied by the data generation process, the probability of necessity equals the reduction factor, and these in turn are related to what epidemiologists call relative risk (RR)

$$PN = 1-(P[y=1|z=1]/P[y=1|z=0]) = 1-RR = RF \qquad (3)$$

(Pearl 2000, p. 292).

In our study, simulating the effect of the proposed speed limit was done by adding a variable z to the pedestrian collision model, with z=0 denoting the current speed limit condition, while z=1 denoted the presence of a strictly enforced 25 mph (40 km/h) limit. The effect of the new limit was simulated using the structural equation

$$v_1 = \begin{cases} v_d, \text{ if } z=0, \text{ or } v_d < v_L \\ v_L, \text{ if } z=1 \text{ and } v_d \geq v_L \end{cases} \qquad (4)$$

where v_L denotes the speed limit, and v_d is the driver's desired speed, assumed to be distributed according to the data collected in the field. In words, equation (4) assumes that when z=1 drivers will travel at their desired speeds as long as these are lower than the speed limit, and at the limit speed otherwise. Probabilities of necessity were then computed for each site using Balke and Pearl's (1994) Twin Network method, in which one first conditions on y=1 and z=0 to obtain posterior distributions for v_1, v_2, x_1, x_2, a, and t_p, and then computes the probability that y=0 using this distribution and v_1 determined according to equation (4). Actual calculations were carried out using Markov Chain Monte Carlo methods, as implemented in the software WinBUGS (Spiegelhalter et al 2000). The results for two sites, 22 and 27b, are displayed in Table 2. Note that the collision probabilities for the case of no speed control, P[y=1|z=0], are the same as the collision probabilities referred to as P[y=1] in the previous section.

At site 22 about 10 simulated pedestrians per 1000 would be struck under the current speed limit condition, and this would drop to 5 per 1000 under the strictly enforced 25 mph (40 km/h) limit. At site 27b, the collision rate would be 64 per 1000 under the current condition, which would drop to 55 per 1000 under the new speed limit. Plugging these conditional probabilities into equation (3) gives reduction factors of RF=1-(.005)/(.01)=0.5 for site 22, and RF=1-(.055)/(.064)=0.14 for site 27b. Note that the reduction factors for the two sites are noticeably different, due to differences in the traffic speeds and volumes. It is this difference that allows us to construct a hypothetical Simpson's paradox. To do this, suppose that an observational study of the effect of the new speed control has been conducted, in which vehicle/pedestrian encounters and collisions were counted, producing the data displayed in Table 3. (Actually, Table 3 was produced by multiplying the probabilities in Table 2 by hypothetical column totals, and then rounding.) Estimating the reduction factors for each site separately gives RF=0.51 (RR=0.49) for site 22, and RF=0.15 (RR=0.85) for site 27b, as expected. Now suppose that instead of analyzing the data for each site separately, the data from the two sites are combined, producing the Both Sites Aggregated portion of Table 3. This leads to an estimated relative risk of RR=2.06, and a corresponding reduction factor of -1.06, implying that the new speed limit would make collisions more rather than less frequent. So should policy be based on the aggregated or disaggregated results? That is, should site be treated as a confounding variable?

Pearl has provided a thorough discussion of Simpson's paradox, and the proper interpretation of this hypothetical result depends on the causal processes underlying the data. Figure 3, adapted from Pearl's Figure 6.2, shows two possibilities, where the dependency structures among the model variables are represented by directed acyclic graphs. Figure 3(a) shows the speed limit policy (z) as having a direct effect on the distribution of vehicle speeds (v), as modeled by equation (4), as well as a direct effect on other site characteristics (s). One way this could occur is that the speed limit policy affected the tendency of pedestrians to appear at one or the other of the sites. If something like this were operating, the interpretation of the results in Table 3 would be that although the effect of the new speed limit was to reduce collision risk at both site 22 and 27b, it also caused more pedestrians to appear at the relatively more dangerous site, 27b. The overall effect then was an increase in accident risk, and the Both Sites Aggregated table gives the appropriate estimate. (But since the exogeneity assumption has been violated, the estimate should not be interpreted as a probability of necessity.) Interestingly, a logically similar argument was advanced by Lave and Elias (1994) regarding the relaxation of the NMSL in the 1980's. On the other hand, Figure 3(b) shows the site characteristics as exogenous. In this case the shift of pedestrians from site 22 to 27b would have occurred even if speed limit policy had not changed, and the appropriate interpretation is made by considering each site separately. An estimate of total effect could then be computed as an appropriately weighted average of the individual effects, but not by using the aggregated data.

Conclusion

The issues raised by the above examples are by no means new, even in highway safety research. Smeed (1955) described an analysis where single and multi-vehicle accidents showed different rates when regressed against annual traffic volumes, while Jovanis and Chang (1986) reported that accidents involving different vehicle types showed different sensitivities to environmental factors. These are situations where different types of entities have different regression coefficients, and so aggregation could lead to ecological biases. With regard to before/after analyses, Tanner (1958) noted that simple aggregation of sites with different effects can give biased estimates of the average effect, and developed an unbiased estimator, while Jarrett (1997) explicitly discussed Simpson's paradox in safety evaluation. What is important to note is that determining the proper interpretation of a statistical study may require prior knowledge of the underlying mechanisms generating the data. Statistical modeling alone may not be sufficient.

These problems are at least in part due to the unfortunate fact that a finite data set is rarely sufficient to uniquely identify a model for the process generating the data. The best that can usually be done is to pre-specify a class of possible models and then identify the member of that class which the data best support. If the true model is also a member of that class then standard consistency results will usually imply that, for adequately large sample sizes, the identified model is a good approximation to the true one. If the true model is not a member of the pre-specified class, this guarantee is lost. For example, the data presented in Figure 1 are consistent with at least two underlying models, which we can call a statistical model and a mechanism model. The statistical model assumes that the collision probability at a site is a random outcome from a population whose mean value is a deterministic function of traffic volume and mean speed. That is, in order to simulate the data in Figure 1 using the statistical model one would first compute the expected value for a site using equation (1), and then take a random draw from a population having that mean value. If the data were actually generated by such a process, aggregate regression modeling would then be sufficient, given a large enough sample, to identify the appropriate variables and coefficient values. The mechanism model on the other hand describes the process actually used to generate the data in Figure 1. The outcome of each individual vehicle/pedestrian encounter is modeled deterministically, but the population of inputs to the deterministic model varies from site to site. Averaging over the population then produces, for each site, a more complex relation between relative collision frequency and aggregate measures such as traffic volume or average speed. Although the statistical model is arguably a simpler and more parsimonious model, in this case it was also the wrong model, and the conclusions it implies concerning the relationship between speed and pedestrian risk are different from those implied by the true model.

The above examples were generated using a plausible, but hypothetical, mechanism for vehicle/pedestrian collisions, and so by themselves provide no evidence as to whether or not road accidents are the result of objectively chancy processes. The chance set-up view has been defended as consistent with probability theory (Hauer et al 2002), but perhaps it would be more helpful to start with the observation that probability, as defined by Kolmogorov's axioms, is an abstract, mathematical object in the same way that a continuous function is. The interesting question then is what sorts of things in the real world behave (at least approximately) like probabilities? As Hacking points out, the notion of chance set-ups is compatible with an interpretation of probabilities as dispositional properties, or propensities, but an equally respectable interpretation (at least in some circles) is that probabilities measure coherent degrees of belief (Gillies 2000). If in fact there is an objectively chancy feature to road accidents one would expect that detailed investigation would uncover a random mechanism (such as radioactive decay) or at least a pseudo-random mechanism (such as a chaotic deterministic subsystem) whose effect is large enough to induce randomness at the macroscopic level (Glennon 1997). But in the investigation and reconstruction of an individual accident, while there may be considerable uncertainty about how the accident happened, this uncertainty is more naturally attributed to ignorance of the accident's particulars, which could in principle be removed with more complete information, rather than to the workings of chance. This latter use of probability, as a logic of uncertainty in accident reconstruction, is developed more fully in Davis (2003).

Absent the identification of a physical mechanism producing chance in individual accidents, maintaining that accidents are objectively chancy would require an account of how the epistemic probabilities one encounters in accident reconstruction become the objective propensities assumed in statistical studies. Alternatively, one could simply maintain that the statistical study of accidents deals with phenomena that are fundamentally different from those investigated in accident reconstruction. As an example of the potential dangers of such a separation we might consider research into the causes of cholera described by Turner (1997). In the mid 19th century it was unclear how cholera was transmitted, but using data from an 1848-1849 epidemic in London, William Farr was able to identify an association between the death rate from cholera and elevation above the Thames river. This association was interpreted as supporting the hypothesis that cholera was an airborne infection. John Snow, on the other hand, looked not only at statistical data but also at individual instances of infection or striking non-infection, and concluded that the disease was transmitted via the patients' evacuations. Although, as Freedman (1999b) has pointed out, in the course of his investigations Snow carried out what can be considered a classic example of a well-conducted observational study using aggregated data, Turner's main point is that Snow did not restrict himself to statistical investigations. Rather, he looked at a variety of evidence, and sought the hypothesis that explained all facts, be they aggregate or individual. I would like to suggest that by viewing accidents as objectively chancy, accident researchers could put themselves in positions similar to Farr's, cut off from the background knowledge needed to design and correctly interpret observational studies. This could even lead to the situation Neil Henry warns of: "Statistics is taught as a method, but too often it becomes the theoretical language of those who adopt the method." (1997, p. 66)

In conclusion, it is well-recognized (Hauer 1990; Davis 2000) that rational management of road safety requires causal models that allow one to predict the consequences of design and operational decisions. If, as I suspect, the processes generating road accidents are more usefully treated as mechanisms rather than chance set-ups, then the most useful level for constructing such models will be that of the individual accident, and this modeling will most usefully be done by treating individual accidents as instances of these mechanisms. As noted earlier, this implies that statistical regularities have no independent status, but rather are merely the result of aggregating particular types and frequencies of mechanisms (Hitchcock 2001). Predicting the aggregate causal effect of some countermeasure would then require a "bottom-up" approach, in which one first identifies the relevant accident mechanisms, then determines the causal effect of the countermeasure on each mechanism, and finally predicts the frequencies of these mechanisms at the location of interest. Because identifying accident mechanisms requires more information than is typically available from computerized accident records, "bottom-up" studies will require more effort than standard statistical studies, and so can be more difficult to carry out.

Table 1. Characteristics of Exogenous Variable Distributions Used in Simulating Pedestrian Accident Risk

Variable	Distribution Type	Parameter Source
Vehicle Speed (v_1)	Normal (î , ó)	ì , ó Estimated for Each Site
Vehicle Headway (h)	Lognormal (î , ó)	ì , ó Estimated for Each Site
Vehicle Initial Distance (x_1)	Deterministic	= v_1h
Driver Reaction Time (t_p)	Lognormal (î , ó)	ì , ó Taken from Fambro et al. 1997
Braking Deceleration (a)	Lognormal (î , ó)	ì , ó Taken from Fambro et al. 1997
Pedestrian Speed (v_2)	Normal (î , ó)	ì , ó Estimated for 5th Grade Boys
Pedestrian Initial Distance (x_2)	Uniform (0,b)	b = Site-Specific Setback

Table 2. Conditional Probabilities (P[y | z]) Simulated for Sites 22 and 27b.

		Site 22				**Site 27b**	
		z=				**z=**	
		0	**1**			**0**	**1**
y=	**0**	.990	.995	**y=**	**0**	.936	.945
	1	.010	.005		**1**	.064	.055

Table 3. Outcome of Hypothetical Before/After Study, Illustrating Simpson's Paradox

Site 22

		z=	
		0	**1**
y=	**0**	2168	816
	1	22	4

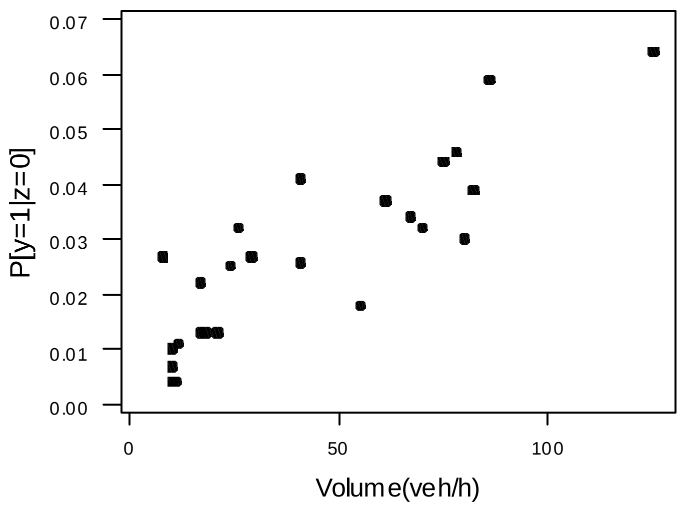
Site 27b

		z=	
		0	**1**
y=	**0**	346	1304
	1	24	76

Both Sites Aggregated

		z=	
		0	**1**
y=	**0**	2514	2120
	1	46	80

Figure 1. Simulated Pedestrian Collision Probabilities versus Traffic Volume and Mean Speed for 25 Streets .



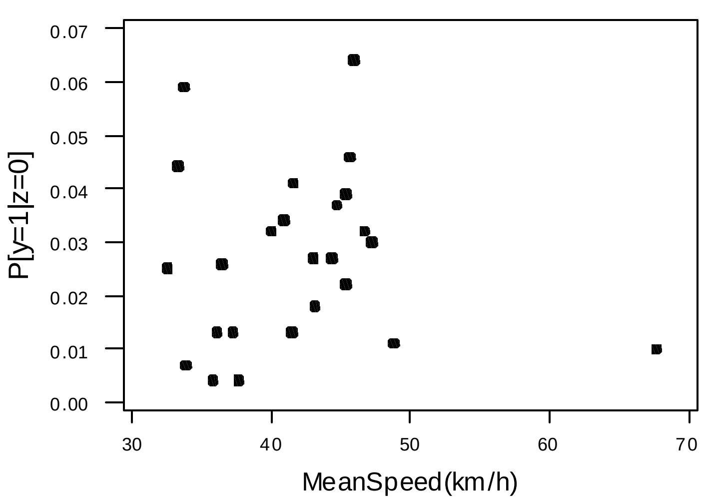


Figure 2. Simulated Relation between Mean Speed and Pedestrian Collision Probability for Two Streets.

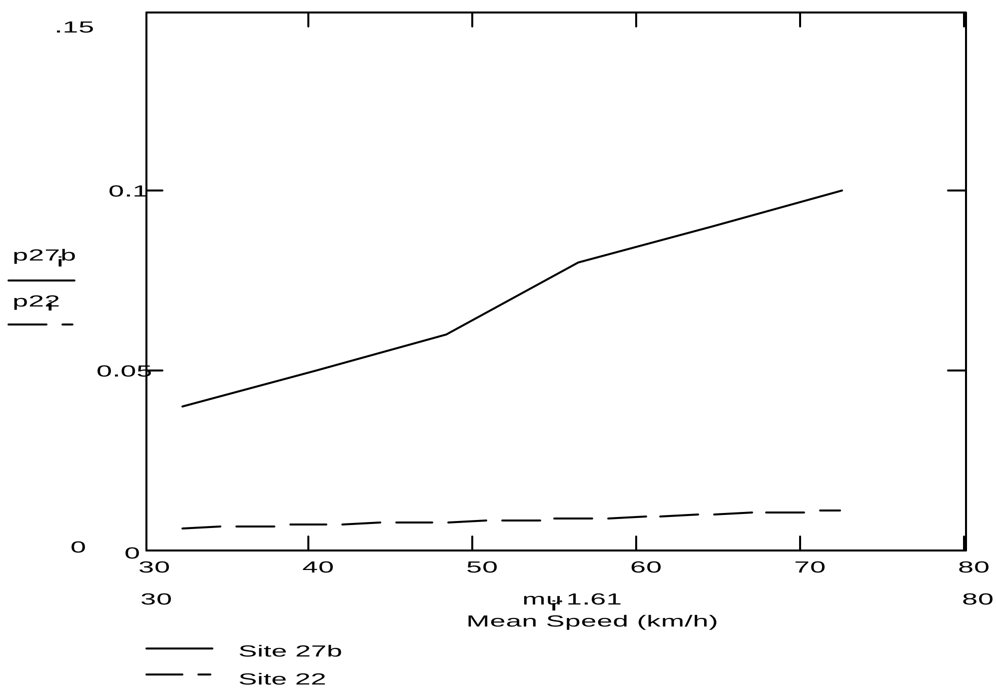
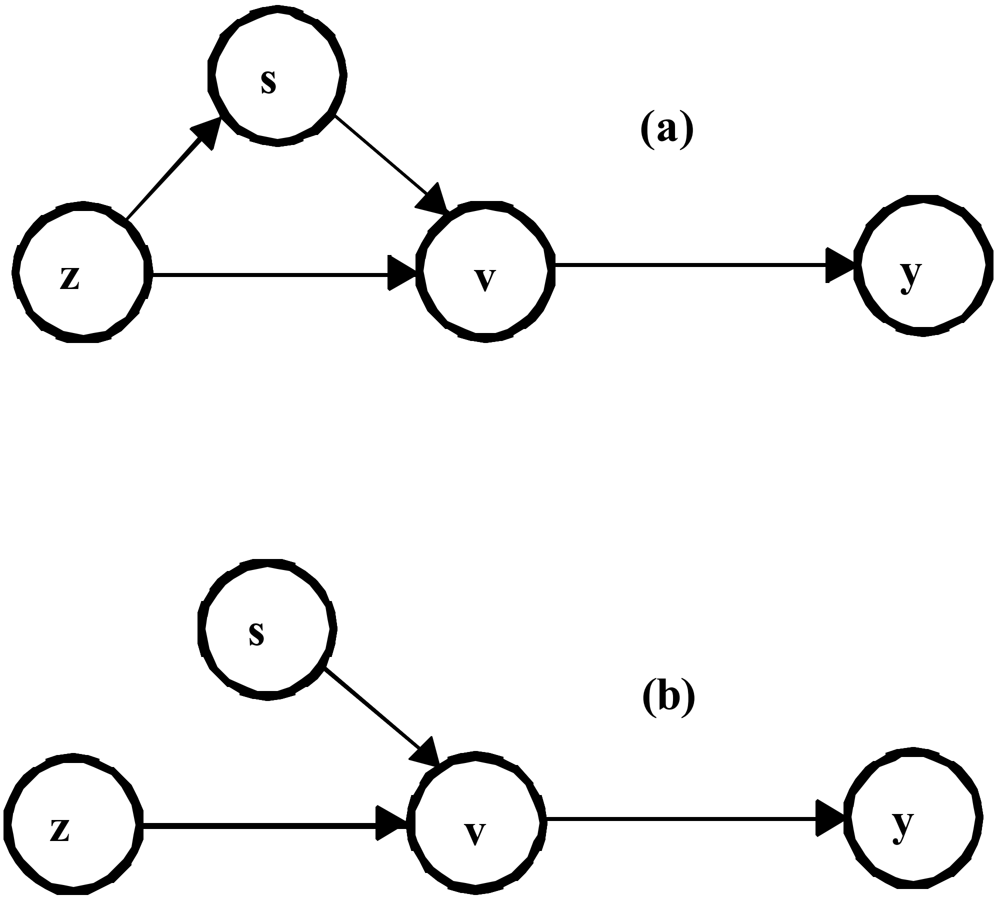


Figure 3. Two Possible Models for the Data in Table 3.



References

Balke, A., and Pearl, J. 1994. Probabilistic evaluation of counterfactual queries. In Proceedings of the 12th National Conference on Artificial Intelligence, MIT Press, Menlo Park, CA, 230-237.

Davis, G., 2000. Accident reduction factors and causal inference in traffic safety studies: A review. Accident Analysis and Prevention 32, 95-109.

Davis, G., 2004. Bayesian reconstruction of traffic accidents and the causal effect of speed in intersection and pedestrian accidents. Technical Report 2 to ITS Institute, University of Minnesota.

Davis, G., Sanderson, K., and Davuluri, S., 2002. Development and Testing of a Vehicle/Pedestrian Collision Model for Neighborhood Traffic Control. Report to Minnesota Dept. of Transportation, St. Paul, MN.

Dawid, A. P., 2000. Causal inference without counterfactuals. J. of American Statistical Association 95, 407-423.

Fambro, D., Fitzpatrick, K., and Koppa, R., 1997. Determination of Stopping Sight Distances. NCHRP Report 400, National Academy Press, Washington, DC.

Freedman, D., 1997. From association to causation via regression. In V. McKim and S. Turner (eds.) Causality in Crisis, University of Notre Dame Press, Notre Dame, IN, 113-162.

Freedman, D., 1999a. Ecological inference and the ecological fallacy, Technical Report 549, Department of Statistics, University of California, Berkeley, CA.

Freedman, D., 1999b From association to causation: Some remarks from the history of statistics. Statistical Science, 14, 243-258.

Garber, N., and Gadiraju R., 1988. Speed Dispersion and Its Influence on Accidents. AAA Foundation for Highway Safety.

Giere, R., 1980. Causal systems and statistical hypotheses. In L. Cohen and M. Hesse (eds.) Applications of Inductive Logic. Oxford University Press, Oxford, UK.

Gillies, D., 2000. Philosophical Theories of Probability. Routledge, London.

Glennon, S., 1996. Mechanisms and the nature of causation. Erkenntnis 44, 49-71.

Glennon, S., 1997. Probable causes and the distinction between subjective and objective chance. Nous 31, 496-519.

Hacking, I., 1964. On the foundations of statistics. British Journal for the Philosophy of Science 15, 1-26.

Hacking, I., 1965. Logic of Statistical Inference. Cambridge University Press, Cambridge, UK.

Hauer, E., 1980. Bias-by-selection: Overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment. Accident Analysis and Prevention 12, 113-117.

Hauer, E., 1982. Traffic conflicts and exposure. Accident Analysis and Prevention 14, 359-364.

Hauer, E., 1990. The engineering of safety and the safety of engineering. In J. Rothe (ed.) Challenging the Old Order: Towards New Directions in Traffic Safety Theory, Transaction Publishers, London, 29-71.

Hauer, E., Kononov, J., Allery, B., and Griffith, M., 2002. Screening the road network for sites with promise. Transportation Research Record, 1784, 27-32.

Henry, N., (1997). Thoughts on the concept and application of statistical models. In B. Van Fraassen (ed.) Topics in the Foundation of Statistics, Kluwer, Dordrecht, 63-67.

Hitchcock, C., 2001. Causal generalizations and good advice, The Monist 84, 218-241.

Holland, P., 1986. Statistics and causal inference. J. of American Statistical Association 81, 945-960.

Jarrett, D., 1997. Assessing the safety effect of treatment using data from a number of sites. Proceedings of Traffic Safety on Two Continents Conference, Swedish National Road and Transport Institute,

Linkoping, Sweden, 145-162.
Jovanis, P., and Chang, H., 1986. Modeling the relationship of accidents to miles traveled. Transportation Research Record 1068, 42-51.
Kapustin, R., 2001. Transportation disaster investigation. In K. Carper (ed.) Forensic Engineering, 2nd edition, CRC Press, Boca Raton, 159-222.
Kloeden, C., McLean, A., Moore, V., and Ponte, G., 1997. Traveling Speed and the Risk of Crash Involvement. NHMRC Road Accident Research Unit, University of Adelaide, Adelaide, Australia.
Lave, C., 1985. Speeding, coordination, and the 55 mph limit. American Economic Review 75, 1159-1164.
Lave, C., and Elias, P. 1994. Did the 65 mph speed limit save lives? Accident Analysis and Prevention 26, 49-62.
Lord, D., Washington, S., and Ivan, J. 2004. Statistical challenges with modeling motor vehicle crashes: Understanding the implication of alternative approaches. Proceedings of the 83rd Annual Meeting of the Transportation Research Board, Paper 04-3162, TRB, Washington, DC.
Lyon, C., Oh, J., Persaud, B., Washington, S., and Bared, J. 2003. Empirical and theoretical support for the IHSDM accident prediction algorithm for at-grade intersection. Paper presented at 82nd annual meeting of Transportation Research Board, Washington, DC.
McCarthy, P., 1998. Effect of speed limits on speed distributions and highway safety: A survey of the literature. In Managing Speed, Special Report 254, Transportation Research Board, Washington, DC, 277-358.
Pearl, J., 2000. Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge, UK.
Persaud, B., Lord, D., and Palmisano, J., 2002. Calibration and transferability of accident prediction models for urban intersections. Transportation Research Record 1784, 57-64.
Robinson, W., 1950. Ecological correlations and the behavior of individuals. American Sociological Review 15, 351-357.
Rohatgi, V., 1976. An Introduction to Probability Theory and Mathematical Statistics. Wiley, New York.
Rubin, D., 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. J. of Educational Psychology 66, 688-701.
Rubin, D., 1978. Bayesian inference for causal effects: The role of randomization. Annals of Statistics 6, 34-58.
Schaffner, K., Clinical trials and causation: Bayesian perspective. Statistics in Medicine 12, 1477-1494.
Shinar, D.,1998. Speed and crashes: A controversial topic and an elusive relationship. In Managing Speed, Special Report 254, Transportation Research Board, Washington, DC, 221-276.
Simpson, E., 1951. The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society, Ser. B, 13, 238-241.
Smeed, R., 1955. Accident rates. International Road Safety and Traffic Review 3, 30-40.
Spiegelhalter, D., Tomas, A., and Best, N. 2000. WinBUGS Version 1.3 User Manual. MRC Biostatistics Unit, Cambridge University, Cambridge, UK.
Tanner, J., 1958. A problem in the combination of accident frequencies. Biometrika 45, 331-342.
Turner, S., 1997. "Net effects": A short history. In V. McKim and S. Turner (eds.) Causality in Crisis. University of Notre Dame Press, Nortre Dame, IN, 23-46.