Vehicle Safety Standard Update: *A Case Study in a Regulatory Debate Using Statistical Models*

Norma Faris Hubele and Mark W. Arndt

In the fall of 2007, the National Highway Transportation Safety Administration (NHTSA), an agency within the U.S. Department of Transportation, issued a proposal to change the safety standard for roof crush of new vehicles under 6,000 lbs. to be sold in the United States. This standard is called the Federal Motor Vehicle Safety Standard 216 (or briefly FMVSS 216). As with all federally proposed changes in vehicle standards, this proposal was posted on the website *www.regulations.gov* in order to solicit feedback from the general public. An important part of NHTSA's rationale for the proposed change relied upon a set of statistical models built using actual rollover crashes experienced across the U.S. The statistical models described severe head, neck and face injuries due to roof contact during rollover crashes. The federal researchers used the statistical models to present their case that roof intrusion increased the risk of occupant injuries and death. They argued that stronger vehicle roofs would maintain occupant survival space during a rollover and consequently improve occupant safety.

WHEN IS A ROLLOVER A ROLLOVER?

It's a rollover crash if, during any portion of the crash, a vehicle rotates 90 degrees or more about its longitudinal or lateral axis. The longitudinal rollover is commonly called the end-to-end roll. The lateral roll is when the vehicle initially rolls to the passenger or driver side.

Three General Stages of Rollover Crashes

1. During the pre-trip phase there is the beginning of the loss of vehicle control. When the trailing wheels lift then the crash has entered phase 2, the trip phase.

2. At the end of the trip phase, the leading wheels lift and the vehicle starts to roll.

3. When the vehicle is actually rolling is considered the third phase. While it is rolling, the vehicle may become airborne, may hit other vehicles or objects, including the ground. The phase ends when the vehicle stops.

FMVSS 216 is meant to address Stage 3, the rolling phase.

As with any federally proposed regulations, there were parties in support of the proposed increase in roof strength and those that opposed the change. But what was particularly interesting about this debate was the fact that alternative statistical models and alternative crash data sets were used to counter and support the federal proposal. We found this to be an excellent example of how different statistical models, using different philosophical approaches and data sets could be used to fuel a debate about how our government would regulate an important industry. The debate took place over a span of about 16 months and was open to the public on documents submitted to the government website. Here, we attempt to step back and examine this debate in the context of vehicle rollover crash analysis and testing research. In a linear narrative we then briefly describe the findings together with the statistical and practical arguments that ensued among the parties. Finally, we discuss the common and competing perspectives on public policy that may motivate the principal participants to help explain the differing approaches. In the end, NHTSA defended their statistical models' results and strengthened the roof test standard, but it was and still is, not without controversy.

The Rollover Crash

Over the past decade, more than one-third of all deaths in cars and light trucks have occurred in rollover crashes. Herein lies the reason that preventing rollovers from occurring and protecting occupants from injury during a rollover has become a central focus of transportation professionals.

The classic examination of motor vehicle injury causation follows the so-called Haddon matrix wherein an examination of the mechanisms causing harm are examined in the temporal phases of the pre-crash, crash and post crash for factors associated with the human, vehicle and environment. The FMVSS are vehicle standards—meaning FMVSS do not set out to regulate human behavior or attributes (for example drunk driving or speeding) or regulate driving environment factors (for example road design or maintenance). Further, FMVSS 216 is a crash phase standard—meaning it is not concerned with pre-crash issues like braking or post crash issues like escapeworthiness or fire.

A common definition of vehicle rollovers is briefly described in the sidebar. The lateral rollover, when the vehicle rolls on its side, is the most common. The longitudinal is the end-to-end roll. The sidebar also gives the phases of the rollover event. The reader may have already guessed that electronic stability control technologies are designed to assist the driver during the first and second phase of the rollover and consequently prevent or reduce the severity of phase three. In this discussion, we are concerned with roof strength, which may play a role in the vehicle protecting occupants during phase three of the rollover event.

We know that a rollover crash is influenced by many factors including: the type of vehicle, driver initiated pre-rollover maneuvers, the roll initiation mechanisms (e.g. surface conditions), the roll location, the roll direction, the number of revolutions of the vehicle and how far the vehicle has traveled during this third phase of the rollover event.

Overall, reduction in occupant exposure to injury is the underlying objective in motor vehicle safety. The fundamental research question that NHTSA was interested in addressing was: taking into account the factors that we know influence occupant injuries in rollover crashes, can we determine a relationship between roof strength and increased risk of occupant injury? In other words, if we increase vehicle roof strength, then would fewer occupants be killed or severely injured in rollover crashes?

In understanding NHTSA's focus on vehicle roof crush it is important to remember that FMVSS216 regulates the vehicle design's roof crush performance in the rolling phase of a rollover crash. While the proposal by the federal analysts was to increase vehicle roof strength requirements, the regulators were fully cognizant that other factors, such as: occupant age, gender; size, seating location, restraint usage, other occupant demographic and pre and post crash environmental factors, effect the potential harm in rollover crashes. These other factors, while important in understanding the syndrome and designing countermeasures of motor vehicle rollover crash injury, are not appropriately addressed by a roof crush standard, but rather by a holistic approach to public policy that encompasses all crash phases and factors.

Rollover and Roof Strength Testing

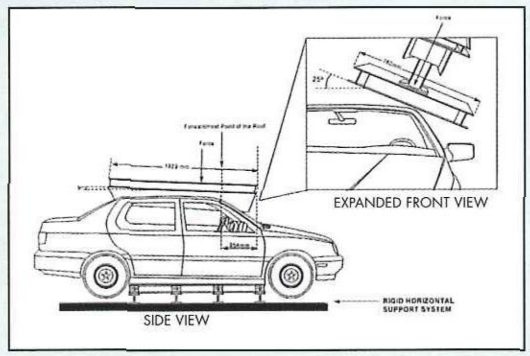
The statistical debate that is the focus of this article is better understood in the context of rollover and roof strength tests together with theories about occupant head, neck and face injuries. Rollover tests were among the first full scale motor vehicle impact test ever performed by the automotive industry and noted to have taken place at the General Motors (GM) Proving Ground as early as 1934. The early tests consisted of tipping a car on its side and letting it roll down a hill. Other early tests consisted of driving cars into a spiral ramp at the top of a hill or driving a vehicle into a sod field inducing ground level rollovers. Unfortunately these early attempts at a rollover test did not produce either repeatable or realistic results. In the early 1970's a dolly rollover test procedure was embodied in the FMVSS 208 regulating vehicle occupant protection systems. The 208 dolly test procedure launches a vehicle 90 degrees to the direction of travel from an initial roll angle of 23 degrees with the leading tires resting against a four inch bar nine inches above the test surface.

In addition to the dolly rollover test procedure, NHTSA was enforcing the FMVSS 216, the roof strength standard (see sidebar for details). This static test was intended to ensure that during the rollover event the vehicle roof would maintain the integrity of the occupant compartment.

FEDERAL MOTOR VEHICLE SAFETY STANDARD (FMVSS)

216 ROOF CRUSH RESISTANCE

Since September 1973, FMVSS 216 has required minimum strength requirements for roof structures over front occupants of passenger cars with Gross Vehicle Weight Rating of 6000 pounds or less. The test procedure involves securing a vehicle on a rigid horizontal surface, placing a flat steel rectangular plate on the vehicle's roof, and using the plate to apply a force of 1.5 times the unloaded weight of the vehicle (up to a maximum of 22,240 N, or 5,000 pounds, for passenger cars) onto the roof structure. A roof must prevent five inches of deformation from the rigid plate as it is quasi-statically forced downward on one side of the roof at a specified orientation and position. Deformation is the measured travel of the rigid plate.



The strength-to-weight ratio (SWR) is the peak strength of the roof obtained when performing this test divided by the vehicle weight.

Strength-to-weight ratio = (peak roof strength/vehicle weight)

The **updated standard** includes a new test procedure effective September 2012 and phased-in through 2016. The updated test procedure requires testing of both the passenger and drivers side of the front roof up to 3.0 times the vehicles unloaded weight prior to interior contact with a presentation of head position or 5 inches of rigid plate travel whichever occurs first. Further the new standard is expanded to include vehicles with a GVWR between 6000 pounds and 10,000 pounds, but at a force of 1.5 times the unloaded vehicle weight.

NATIONAL AUTOMOTIVE SAMPLING SYSTEM CRASHWORTHINESS DATA SYSTEM

(NASS CDS)

The NASS CDS database consists of detailed information about crash events, the vehicles, the occupants, and the associated injuries for approximately 5,000 crashes statistically sampled annually. Crashes are assigned weights inversely proportional to their sample probabilities; as a result, large weights can be assigned to some sampled crashes and small weights to others. Multiple sources of information, including the police accident report, on-site investigations of the crash scene and vehicle, as well as detailed measurements, such as crush profile of the vehicle and occupant compartments, are used to construct each record in the database. In addition, detailed injury data from hospital records and interviews with medical personnel are also reported for each occupant or pedestrian in a crash. For more information see *www.nhtsa.gov/NASS*.

The 208 dolly rollover test was the protocol utilized in the now famous Malibu tests performed by GM and first published by the Society of Automotive Engineers in 1985. Four of the tests used GM Malibu passenger cars with production roofs (i.e., roofs that had passed the FMVSS 216 standard at the time), while four used roll cages; all had unrestrained dummies. Regarding the issue of roof strength, they concluded that roof deformation relative to the seat had no effect on injury mechanism and that stronger roofs did not have any increased level of protection over the standard roof vehicles. Notably, during the rollover, the unrestrained dummies moved upward and outward relative to the vehicle and in instances of injurious neck force the dummy head was against the roof early in an impact before distance between the roof and seat was decreased. The tests did not support the theory that roof crush drives an occupant into the seat. In addition, they noted that partial or complete ejection was a consequence of unrestrained occupants in rollover crashes. Results of the test supported the conclusion that it is generally more desirable to remain contained within the car than to be ejected in a rollover collision.

The Malibu II tests published in 1990 consisted of an identical set of experiments plus inverted static tests and an inverted drop test; here they used restrained dummies. The sidebar on page 7 gives a list of the major findings of these tests. The Malibu studies describe a compressive neck injury mechanism that occurs after head-to-roof contact, but before the roof crushes, and found no relationship between roof crush and neck injury. Pre-roof crush head-to-roof contact was described by way of the dummy's buttock motion off the seat. A summary included: "The absence of deformation may benefit belted occupants if it results in the belted occupant not contacting the roof. The reduction of roof deformation in the rollcaged vehicle had no effect in reducing neck loads for the dummies in the area of ground impact." These seemingly contradictory statements are perhaps a stronger comment on the seatbelt system used to restrain the dummy in the 1980's vintage Malibu vehicle. The seatbelt system unlike available for modern vehicles did not have collision sensing mechanisms that pretension and reduce occupant motions relative to the seat.

NHTSA's Statistical Approach

During the 1980s and 1990s, while these Malibu tests results were being debated, NHTSA launched its biggest and most extensive data collection effort on a sample of crashes nationwide called the National Automotive Sampling System Crashworthiness Data System (NASS CDS, see sidebar for more information). In addition to collecting vehicle measurements, investigators also collect information from medical records and interviews with health professionals treating the injured occupants. Consequently, we have information about the various injuries incurred by the occupant, as well as the source of the injury. In this application, the federal analysts were particularly interested in head, neck and face injuries identified as having resulted from roof contact in rollover crashes.

The NHTSA statistical study discussed here was a follow-up to two past studies, but much more extensive (see Further Reading). Here they defined a specific set of rollover crashes by the criteria listed in Table 1 to assess the occupant benefits of stronger roofs. These criteria, they argued, matched the target population for the FMVSS 216 under consideration. In particular, NHTSA chose to include only single vehicle rollover crashes because they wanted the least complicated events. As is seen in the sidebar, FMVSS 216 applies a rigid plate or platen over the front occupants' compartment; hence NHTSA only selected occupants of front, outboard seats. They required the occupant to be belted by lap and shoulder to minimize the notion that the injury was a result of the occupant "diving" into the roof; this also required that the occupant age exceed 12 years. Since NHTSA is particularly interested in how belted occupants who are contained within the vehicle fare in rollover crashes, they restricted their occupants to non-ejected. The fourth criterion specifies that the vehicle must have rolled at least 180 degrees or at least two quarter turns (four quarter turns would result in the vehicle coming to rest on its wheels). The minimum 2 quarter turns necessitates that the vehicle had the potential to have roof damage due to roof-to-ground contact. More complex rollover events that may have incurred roof damage due to impact with another fixed object, such as a pole or tree, were also eliminated. Another filter that excludes fewer data was the requirement that the vehicle could not be towing a trailing unit nor could it be multistage or certified altered vehicle.

NHTSA selected these occupant records conforming to these criteria from the annual survey sample NASS CDS data for calendar years 1997 through

2005. They used the data to assess if a statistically significant relationship existed between the maximum head, neck and face injuries of rollover occupants due to roof contact and the amount of roof intrusion or post-crash headroom. In other words, after the crashes they studied what happened to the occupant and tried to relate the observed injuries to the observed roof damage of the vehicle. NHTSA carried out an extensive analysis building 24 models. Twelve models used an ordered probit (aka cumulative probit model) since the dependent variable was measured on an ordinal injury scale from 0 representing no injury to head, neck or face due to roof contact, to 6 representing maximum, untreatable (i.e., fatal) injury. The 0 to 6 scale followed coding pursuant to the Abbreviated Injury Scale, a consensus-based threat to life injury scale. (See the sidebar for more explanation of the statistical models.) Another twelve models were binary probit models with 0 representing either no or minor head, neck or face injury due to roof contact, 1 otherwise.

These models were further divided by how they represented roof damage. Some models used the explanatory variable vertical intrusion, others used post-crash headroom. In some models these explanatory variables were expressed in inches, in others, NHTSA used a dichotomous variable. (e.g., there was no headroom or there was some remaining headroom). A headroom variable represents the space between the top of the head and roof. Headroom effect would be influenced by a seatbelt's ability to hold an occupant in the seat and by vehicle design interior height and occupant stature. Different vehicles have different interior height—it's been said that pickup trucks are made so a driver can wear a cowboy hat.

Furthermore, the NHTSA models were divided into two groups, those that were 'adjusted' and those that were not. These adjusted models had additional independent variables, besides the roof damage measurement. They used either the number of quarter rolls or the number of roof exposures to the ground as a dependent variable. The number of roof exposures is merely the number of 'potential' times that the roof may have come in contact with the ground. This number is necessarily a function of the number of quarter rolls and therefore is merely an alternative representation of the severity of the crash. In addition, NHTSA had an indicator variable noting whether or not the vehicle experienced a lateral or longitudinal roll. They used occupant age and gender, as well as another indicator variable noting whether or not the vehicle was a light truck (e.g., SUV, light pickup) or a passenger car. Table 2 on Page 9, shows the list of all the independent variables considered to be potentially confounding factors in NHTSA's adjusted models.

All these models arrived at the same conclusion:

In all 24 models, the relationship between injury severity and the explanatory variable, intrusion or headroom, was statistically significant.

Table 1. Criteria Used to Select Occupant Crash Records from the NASS CDS for NHTSA 2007 Statistical Study

1. Single rollover vehicle crashes only
2. Occupants 13 years old and older
3. Front, outboard-seated occupants wearing a lap and shoulder belt who were not ejected
4. Rollover crashes in which at least 2 quarter turns occurred
5. Non-arrested rollovers (e.g., the vehicle did not come to rest against a pole)

1990 GENERAL MOTORS MALIBU II TESTS USING RESTRAINED DUMMIES

SUMMARY FINDINGS REGARDING DUMMY KINEMATICS AND NECK INJURY MECHANISM

1. A belted 50th percentile male has his head positioned against the roof when hanging statically upside down.

2. Neck loads resulted from "diving" type impacts, in which the head stops and the torso momentum compresses the neck.

3. Roof deformation never caused the dummy to be compressed between the roof and seat.

4. Safety belts prevented ejection and projected impacts with the vehicle interior.

5. Safety belts did not result in reduced head/neck loads for dummies in the area of a ground contact.

6. Under similar impacts, there was no increased protection in rollcaged vehicles over the production roof vehicles.

7. The absence of deformation may benefit belted occupants if it results in not contacting the roof.

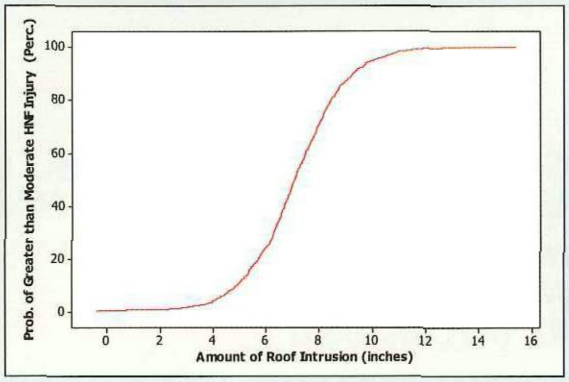
8. Reduction of roof deformation in the rollcaged vehicle had no effect in reducing neck loads for the dummies in the area of ground impact.

Malibu II suggested improved rollover crash protection, particularly reduced roof deformation, could mitigate axial neck loads if safety belt systems prevented substantial occupant motion that keeps the head away from the area of impact.

STATISTICAL MODELS

The models used by the researchers in this case study fall under the heading of generalized linear models. All analysts built models for binary responses. For example, IIHS, JP Research and Biomech, as well as Wecker, coded each occupant rollover crash record with a response of killed/ incapacitated (1) or not (0). In these binary response type models, the analysts assume that the relationship between the explanatory variables and the probability of outcome resembles an S-curve such as shown below. In this illustration based on the NHTSA study, the assumption is that the probability of a moderate or worse head, neck or face injury *increases non-linearly with the amount of roof intrusion.*

Illustration of the Assumed Relationship between the Response Variable and Explanatory Variable Using Fictitious Data (HNF represents Head, Neck or Face)



The S-curve resembles a cumulative distribution function and conforms to the requirement that the probability lies between 0 and 1 (or 0 and 100 when the probability is expressed as a percent). The logistic distribution has such a cumulative function. Consider

$$\pi(x) = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$$

Then the transformation

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right]$$

yields the relationship

$$g(x) = \alpha + \beta x$$

The transformation function $g(x)$ is called the logit link or logit transformation. It has the desirable properties of being linear in the parameters and being able to take on values in $(-\infty, \infty)$. When the binary coded data, together with the explanatory variables are input into a software system, the user is given an opportunity to specify the choice of cumulative distribution function to fit the S-curve or, equivalently, the link function. IIHS, JP Research and Biomech, and Wecker chose the logistic distribution resulting in a logistic regression model. NHTSA chose the normal distribution function resulting in a probit model. The logistic distribution is symmetric and bell-shaped but has slightly thicker tails than the normal distribution. NHTSA also used a cumulative probit model to model the probabilities for the increasing injury levels. More information can be found in texts such as Agresti's sited in Further Reading.

Feedback and Dispute

Some opponents to the NHTSA findings argued that the statistical relationship that NHTSA established in their models did not imply causation between roof intrusion and injury. NHTSA agreed with the general principle, however, they countered that injuries used in their models were specific to injuries in which the roof was the injury source as recorded in the NASS CDS. Consequently this statistical relationship came close to the notion of establishing a causal link. Skepticism regarding a causal relationship between roof intrusion and injury may be rooted in the experimental results of the Malibu studies.

In response to criticisms that not all rollovers are the same, NHTSA argued that their models controlled crash severity using number of quarter rolls or number of potential roof-to-ground contacts and its effect on injuries. Finally, NHTSA claimed that they had reduced the effect of the occupant "diving" during the rollover event by including only belted occupants.

The first NGO statistical study submitted to the docket was performed by the Insurance Institute for Highway Safety (IIHS). IIHS's logistic models estimated the risk of a killed or incapacitated driver injury in a single vehicle rollover crash. Unlike the NHTSA study that relied totally on detailed crash data, IIHS data was a combination of rollover test roof strength measurements of 11 midsize four-door SUV models (and their corporate twins, think Ford Explorer and Mercury Mountaineer; same underlying vehicle, but different cosmetic look and name) and their historical rollover crashes as reported in police accident reports from 12 states. The roof strength measurements were obtained from a combination of NHTSA test data and IIHS contracted testing using the FMVSS 216 procedure. Among the test measurements analyzed by IIHS was the peak strength-to-weight ratio within 5 inches of roof crush, as defined in the sidebar on FMVSS 216. This test information was combined with crash information from Florida, Georgia, Illinois, Kentucky, Maryland, Missouri, New Mexico, North Carolina, Ohio, Pennsylvania, Wisconsin, and Wyoming. These states use the same KABCO injury coding system, where K indicates a fatal injury, A incapacitating and BCO lesser injuries. In their logistic model (again see the sidebar on statistical models), IIHS grouped KA versus BCO (as a binary response variable) to define the driver's injury level.

In direct comparison to the NHTSA data, this driver injury information is not specific to any specific body part, i.e., not necessarily head, neck or face; nor is it known if the injury is due to roof contact. With regard to the NHTSA criteria listed in Table 1, IIHS chose only drivers and single vehicle rollover crashes. Because the state data only shows if the vehicle rolled, and not the number of quarter rolls incurred, IIHS was not able to limit their data to rollovers with at least 2 quarter rolls. About 25% of SUV rollovers are estimated to have only 1 quarter roll (and therefore would not have the roof damage of interest here). Nor could IIHS eliminate arrested crashes—rollovers that stopped against an object like a barrier, tree or pole.

In their final statistical models, IIHS controlled (or "adjusted" using NHTSA's terminology) for driver age, static stability factor and individual state difference. They explained their use of the static stability factor (i.e., the height of the vehicle's center of gravity divided by two times the track width) because less stable vehicles could roll more easily in low speed, less severe crashes. Static stability factor is used by NHTSA to rate vehicle rollover propensity in their 5-star rating system.

Table 2 shows other explanatory variables that were investigated by IIHS, but found to be insignificant (NS). IIHS argued that belt use coded in the state databases is unreliable and may be correlated to injury status. For example, a coherent non-incapacitated driver could falsely claim proper belt use. To address this issue, IIHS built multiple models. They concluded that models including all driver and all belted drivers separately arrived at the same conclusion (i.e., increased roof strength was statistically linked to reduced risk of driver incapacitating injury or death). They found no statistical link between increased roof strength and unbelted drivers; in this phase IIHS did not investigate ejection/non-ejection.

After IIHS submitted their report, JP Research and Biomech, two consulting firms, jointly published an 11-page counter-report. Using the same data as IIHS and the same type of logistic models, JP Research and Biomech argued for more control variables in their models, as shown in Table 2. They stated that the aspect ratio should be included, and not the static stability factor, in their models. Whereas static stability factor is related to propensity to roll, the aspect ratio, which is the ratio of the vehicle height to track width, gives a measure of vehicle profile. They found this factor to be significant.

In the state databases alcohol is coded as either the blood alcohol content of the driver and/or the general statement that driver "had been drinking." JP Research and Biomech stated that this variable should be included based on their own previous investigations. Here, they found that alcohol, using either coding procedure, was statistically significant in predicting driver injury level.

JP Research and Biomech also found that when an urban/rural indicator of the location of the crash was available, the predictive models performed better. Again relying on their own previous work, JP Research and Biomech included vehicle weight in their models and found it to be statistically significant. Citing the summary statistics that 56% of the fatalities and 28% of the serious/fatal injuries were unbelted and completely ejected, JP Research and Biomech argued that both the belt status and ejection status of the driver should be included in the models. They found both of these explanatory variables to be significant.

All the models investigated by JP Research and Biomech, including those for fatal injuries only or fatal injuries combined with serious injuries,

Table 2. List of Independent Variables Used in the Different Statistical Models; an X Indicates This Explanatory Variable Was Investigated by at Least One of the Models, and a Blank Indicates It Was Not Included in Any Published Model

	NHTSA	IIHS, Wecker	Padmanaban & Moffatt
Response Variable			
Head, Neck or Face Injury Due to Roof Contact	X		
Killed or Incapacitating Injury		X	X
Explanatory Variables			
Crash Information			
Post-Crash Headroom	X		
# Quarter Turns	X		
Longitudinal or Lateral Roll	X		
Alcohol			X
Urban vs. Rural Location			X
Vehicle Information			
Type of Vehicle	X	11 SUV's (roof designs)	11 SUV's (roof designs)
Drive Type (two vs. four wheel)		X-NS	X-NS
Weight			X
Age		X-NS	X-NS
Static Stability Factor		X-NS(a)	
Aspect Ratio			X
Roof Strength-to-Weight Ratio		X	X-NS
Occupant/Driver			
Age	X	X	X
Gender	X	X-NS	X
Belt Status	(Only Belted)	X(b)	X
Ejection Status	(Only Non-ejected)	X(b)	X
State			
Indicator		X	X

Notes: (a) Although IIHS found this factor to not be statistically significant at the 0.05 level, they still included it in their final model.

(b) IIHS built various models.

found that the SWR roof strength measurement at the 5 inches as performed using FMVSS 216 was not a statistically significant factor in driver injuries. They found the same result for unbelted, non-ejected drivers.

Finally, JP Research and Biomech pointed out that while that data included about 23,000 crashes, only four models accounted for nearly 75% of the data in the IIHS database. Consequently, these models dominated all the analyses.

About 6 weeks later IIHS responded. They provided evidence that for their data set ejection was related to roof crush and therefore defended their not including ejection status as an explanatory variable in their initial study. They performed their study for non-ejected and again concluded that roof strength was statistically significant in predicting injuries to these drivers. IIHS continued to defend their exclusion of belt status, citing other statistical studies that called into question the reliability of the police reported belt use.

Regarding alcohol, IIHS rebutted the work of JP Research and Biomech by citing the low frequency of actual BAC information in the crash records, inconsistencies in the data and assumptions used to investigate this factor. IIHS also questioned JP Research and Biomech's statistical interpretation of significance levels and cut-off values. They defended their use of static stability factor and criticized, as unjustified, use of the aspect ratio. They also pointed out that vehicle weight was already included in the roof strength measurement SWR as the denominator and called into question why it was needed as a separate explanatory variable.

Since JP Research and Biomech used more independent variables in their models and since some state data crash records were missing this additional information, then JP Research and Biomech had to use a smaller data set to build their models. IIHS stated this as evidence that the two antagonistic approaches were actually performed on two different sets of data, and JP Research and Biomech did not in fact 'duplicate' the IIHS studies.

About 2 weeks after the IIHS response to the JP Research and Biomech criticism, JP Research and Biomech again submitted a study to the docket. Here they provided more details and references supporting their position. Again they defended their modeling approach that more explanatory variables should be included in the models. Regarding ejection, they used statistical models to show no relationship to roof strength but a statistically significant relationship to belt use, alcohol and the rural/urban indicator. Regarding belt use, they performed simulation studies that defended their original use of the police coded belt status. Regarding alcohol, they cited studies that indicated that high BAC is related to increased injury severity. Regarding the use of an urban/rural indicator, they repeated their findings that including this variable in the models resulted in roof strength becoming insignificant. Regarding the use of aspect ratio instead of static stability factor, JP Research and Biomech again repeated their opinion that this was a more logical vehicle characteristic for investigating occupant protection in a rollover. They provided detailed statistical output documenting their models.

Critique of the IIHS study continued from the Alliance of Automobile Manufacturers in a report by Wecker and Associates. Unlike the earlier criticisms, Wecker and Associates did not spend any time arguing the merits of the models. Rather, they adopted the IIHS models but changed the data. In their 12-page report, Wecker and Associates documented an analysis, using data from 9 of the 12 states used by IIHS (Wecker and Associates were unsuccessful in getting data from Georgia, New Mexico, and Kentucky).

Wecker and Associates argued that the IIHS assumption of a single statistical model for all the roof strengths did not hold up. Their arguments were based on several models using smaller data sets for different intervals of vehicle roof strengths—some weaker roofs and some stronger roofs. They presented evidence that when these IIHS-type models were applied to the vehicles with the stronger roofs of at least 2.0 SWR values, there was no statistical relationship between injury risk and roof strength. Using this research, the Alliance of Automobile Manufacturers went before a congressional committee and testified, "we do not see a causal relationship with injury risk and roof strength."

But IIHS did not sit quietly. In June 2008, they issued a rebuttal to the alliance's claim. IIHS mimicked the study performed by Wecker and Associates. First, IIHS reduced their 12-state database to the nine states used by Wecker and Associates and showed that the data did not line up. Wecker and Associates nine-state data had higher rates of incapacitating injuries and fatalities, more records for 2-wheel drive vehicles and few records for 4-wheel drive vehicles. This was a mystery that did not get resolved.

Next, using their original 12-state database, IIHS performed a sensitivity analysis on the models built by Wecker and Associates on the vehicles with the higher roof strengths of at least 2.0 SWR. They found that with cut-off values of 1.9, 2.0 or 2.1 they still showed reduced injury and fatality risk as the roof strength increased. IIHS also investigated the fatality risk only for data with SWR values above 2.0 and 2.5. Again, they upheld their findings that increased roof strength decreased the risk of fatality. In the end, IIHS affirmed its original recommendation that FMVSS 216 should increase the minimum force requirement to a value beyond 2.5, as originally proposed by NHTSA.

Shortly before NHTSA published their final ruling on the proposed new standard, IIHS published another study, similar to the SUV study, but this time they had 12 smaller passenger cars. They addressed some of JP Research and Biomech's criticisms by including a separate model for non-ejected drivers and one for belted drivers. IIHS refuted the claim that a rural/urban covariate eliminated roof strength as a significant factor. This additional research supported their earlier finding: stronger roofs offered drivers better protection.

In their final ruling on the proposed standard, NHTSA addressed this "boxing match"—as one safety organization called the debate. NHTSA used the term "worrisome" when referring to the exclusion of belt status in the IIHS models. They also focused on the fact that only four models accounted for the high proportion of the data (nearly 75%), speculating that some of the factors that JP Research and Biomech found to be insignificant (or significant) may have been a result of the lack of variability in the roof strength measurement (since nearly 75% of the observations had the same roof strength measurement). In addition, the inference issue was raised by NHTSA. Do these limited number of SUV models provide sufficient evidence for the rest of the fleet of vehicles on the road?

When NHTSA reviewed IIHS later work, they again criticized IIHS since only three passenger models accounted for the majority of the observations used in the study. This again raised the specter that the statistical results were a consequence of these specific cars and did not necessarily provide robust, statistically valid results for other vehicles.

IIHS repeatedly argued that NHTSA underestimated the reducing injury effect of stronger roofs. NHTSA responded that IIHS was overestimating the effect due to their different modeling approaches regarding driver safety belt status, treatment of ejections and the impact of a handful of models on their results.

In the end, NHTSA defended its approach which focused solely on head, neck and face injuries due to roof contact, which used roof intrusion as a model input and which excluded one-quarter roll vehicles from their analysis. In summary, their models more closely addressed the critical research question: would increasing roof strength reduce the risk of head, neck and face injuries due to roof contact?

In one final reflection on the controversy about how to handle ejected drivers, NHTSA performed additional research and found consistent with all others that the magnitude of the roof intrusion was unrelated to whether or not the occupant was ejected.

Discussion

The authors of this article do not speak for the various participants in the roof crush debate. However, with a presumption that all acted in the interest of improved motor vehicle crash safety, a discussion on their public policy perspectives may be useful in understanding the different statistical approaches. Essentially, these participants in the roof crush debate represent three distinct institutions: governmental regulators; automotive manufacturers and insurance companies. Listing these institutions should conjure the economic and political forces at play.

Governmental regulators at NHTSA were first legislatively charged under the National Traffic and Motor Vehicle Safety Act of 1966 with promulgating minimum standards that protects the public against unreasonable risk of accidents and death or injury in an accident. Further, NHTSA is statutorily charged to reduce traffic accidents and deaths and injuries resulting from traffic accidents. FMVSS216, the roof crush standard, specifically applies to vehicle performance in the rollover phase of the Haddon matrix. NHTSA has extensive research programs and standards that cover other variables identified by participants in the roof crush standard debate. Contrasting NHTSA's statistical result with experimental results may suggest that the Malibu studies insufficiently described the injury exposure and/or mechanism in rollovers.

Motor vehicle manufacturers are for-profit institutions that must comply with government safety regulation. On one hand, regulation levels and limits the playing field. All manufacturers have to meet the same standard—so presumably all have a comparable economic cost to comply with a standard and new competitors are limited by the economic hurdle of compliance. On the other hand, manufacturers independently develop their own commitments to safety. For example, a manufacturer's view of rollover safety, influenced by experimental results that do not establish a relationship of roof crush to specific rollover injuries, may see greater efficiencies in addressing parts of the Haddon matrix that are outside its control. The controversies regarding inclusion of variables addressing alcohol, seat belt use and rural versus urban crash location may be a manifestation of a perspective that shifts safety intervention from the vehicle to the driver and/or environment. The public policy of a manufacturer would manifest the most efficient allocation of resources toward the safe use and safe performance of its product.

According to its web site the IIHS is an independent, nonprofit, scientific, and educational organization dedicated to reducing the losses—deaths, injuries, and property damage—from crashes on the nation's highways. In August of 2011 the Institute described itself as being wholly supported by a list of auto insurance companies and three insurance associations. The list reads like a who's who of auto insurers. William Haddon, the namesake of the Haddon matrix and the first Federal Highway Safety chief, was president of the Institute from 1969 until his death in 1985. Like auto manufacturers, many insurance companies are for-profit, however, the auto insurers pay the costs of accident and occupant injury and death. Auto insurers benefit directly in interventions that decrease costs and increase predictability of outcome. To this end, IIHS has led NHTSA and manufactures in developing stringent independent assessments of vehicle performance. IIHS statistical analyses that supported the standard's change and sparred with the manufacturer's reports reflect a public policy position that is consistent with its own newest test assessment of vehicle roof strength.

Closing Comments

The public debate around FMVSS 216 constitutes a classical example of how data and statistical analysts influence our lives. This article provides no real surprises: different analysts using different data sets, different explanatory variables and different statistical models arrive at different conclusions. In the end, the decision maker or the reader is left to decide which point of view more closely aligns with his/her own. NHTSA listened to each of these analysts and made their collective decision. They decided that their original thesis and statistical results were stronger than the opposing points of view.

As stated by NHTSA, "The intent of this rule is to improve occupant protection in rollover crashes that involve roof crush." According to their models, the changes in FMVSS No. 216 will prevent 135 fatalities and 1065 nonfatal injuries, annually. In the future, we can expect to see NHTSA analysts and others again using crash data and statistical models to answer the question: did the increase in roof strength truly save the lives and reduce the injuries predicted? ■

Further Reading

National Highway Traffic Safety Administration. 2008. Federal Motor Vehicle Safety Standard: Roof Crush Resistance. *Docket No. NHTSA-2008-0015 RIN 2127-AG51* National Highway Traffic Safety Administration. Washington, DC.

Strashny, A. 2007. "The Role of Vertical Roof Intrusion and Post-Crash Headroom in Predicting Roof Contact Injuries to the Head, Neck, or Face During FMVSS No. 216 Rollovers; An Updated Analysis." Unpublished but available in *NHTSA dockets NHTSA –2208-0015*. National Highway Traffic Safety Administration. Washington, DC.

National Automotive Sampling System Crashworthiness Data System. *www.nhtsa.dot.gov*

Brumbelow, M. L., Teoh, E. R., Zuby, D. S. And Mccartt, A. T. of the Insurance Institute for Highway Safety. 2008."Roof Strength and Injury Risk in Rollover Crashes." Unpublished but available online as document *NHTSA-2008-0015-0077.1 www.regulations.gov/fdmspublic/component/main?main=DocketDetail&d=NHTSA-2008-0015*

Brumbelow, M. L., Teoh, E. R. 2009. "Roof Strength and Injury Risk in Rollover Crashes of Passenger Cars and SUVs." Presented at the 21st International Technical Conference on the Enhanced Safety of Vehicles, Stuttgart, Germany. *www.iihs.org/research/topics/pdf/r1117.pdf*

Padmanaban, J. and Moffatt, E. 2008. "JP Research's Response to IIHS Docket Submission of May 13, 2008." Unpublished but available as document *NHTSA-2008-0015-0091.1 www.regulations.gov/fdmspublic/component/main?main=DocketDetail&d=NHTSA-2008-0015*

Padmanaban, J. and Moffatt, E. A. 2008. "Evaluation of IIHS Study: Roof Strength and Injury Risk in Rollover Crashes." Unpublished but available online as document *NHTSA-2008-0015-0081.1 www.regulations.gov/fdmspublic/component/main?main=DocketDetail&d=NHTSA-2008-0015*

Agresti, A. 2007. *An Introduction to Categorical Data Analysis*. New York: Wiley.

About the Authors

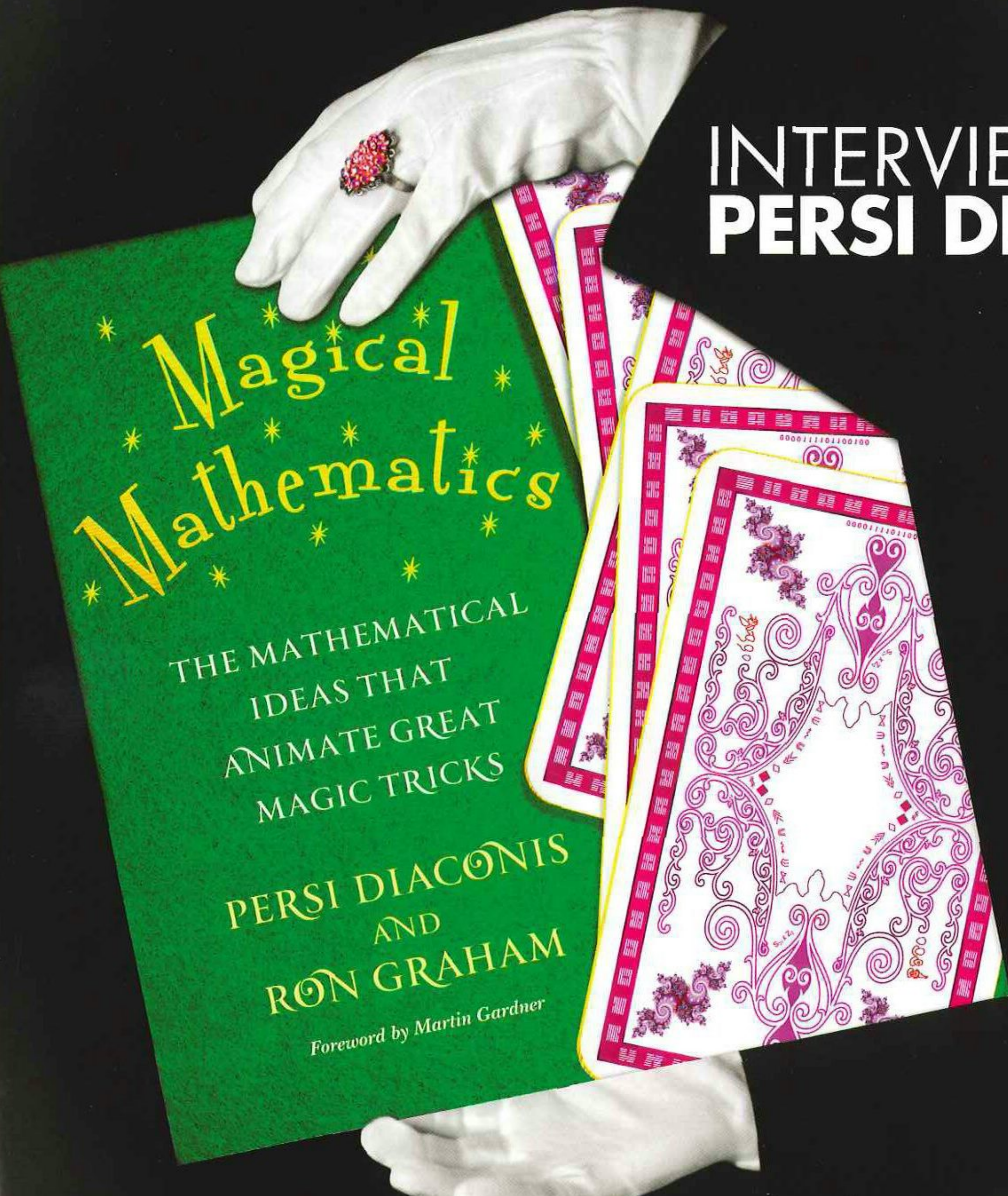
Norma Faris Hubele is professor emeritus in statistics and industrial engineering at Arizona State University. Her current research focuses on automotive safety statistics, in particular, and statistics and the law, in general. She also is co-founder of a metallurgical processing firm, Refrac Systems, in Chandler, Arizona.

Mark W. Arndt is president and founder of Transportation Safety Technologies, Inc. in Mesa, Arizona. Arndt has published research in the field of rollover crashworthiness and consults on engineering aspects of motor vehicle rollover crash and injury causation.

Vol. 25, No. 4, 2012

CHANCE

A Magazine for People Interested in the Analysis of Data



INTERVIEW WITH PERSI DIACONIS

Also...

Ethics and the Statistical Use of Prior Information

Prediction Markets: How Accurate Are They?

09332480(2012)25(4)

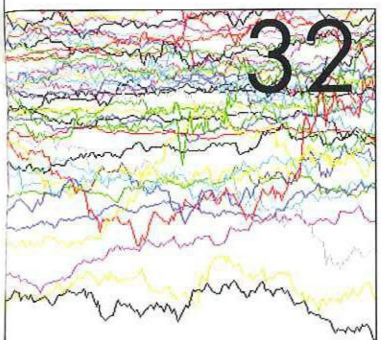
Taylor & Francis
Taylor & Francis Group

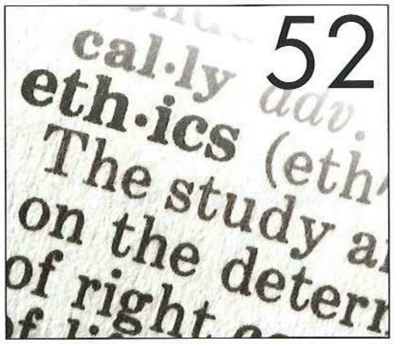
ASA
AMERICAN STATISTICAL ASSOCIATION

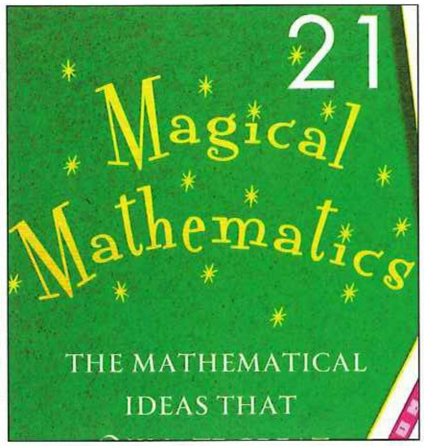
CHANCE

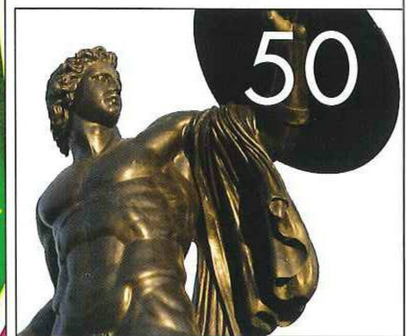
A Magazine for People Interested in the Analysis of Data

http://chance.amstat.org









ARTICLES

COLUMNS

DEPARTMENTS

Abstracted/indexed in Academic OneFile, Academic Search, ASFA, CSA/Proquest, Current Abstracts, Current Index to Statistics, Gale, Google Scholar, MathEDUC, Mathematical Reviews, OCLC, Summon by Serial Solutions, TOC Premier, Zentralblatt Math

Cover image by Kathryn Wright