Traffic Injury Prevention, 3:288–293, 2002

1538-9588/02 $12.00 + .00
DOI: 10.1080/15389580290129044

How Do Euro NCAP Results Correlate with Real-Life Injury Risks? A Paired Comparison Study of Car-to-Car Crashes

ANDERS LIE

Swedish National Road Administration, Borlänge, Sweden

CLAES TINGVALL

Monash University Accident Research Centre, Clayton, Victoria, Australia

The European New Car Assessment Program (Euro NCAP) is a resource for consumers regarding vehicle crash safety. The program promotes safety developments, and credits car manufacturers for focusing on safety. This study was based on real-life car-to-car crashes and results show that the overall indication of the safety level, provided by the Euro NCAP crash testing, is a valid prediction when considering severe or fatal injuries. No significant injury risk differences were seen for minor injury crashes. In car-to-car collisions, cars with three or four stars were found to be approximately 30% safer when compared with two-star cars or cars without a Euro NCAP score. The strong correlation between injury risk and Euro NCAP scores is not necessarily similarly good for individual car models. Pedestrian safety and child occupant protection were not considered in the present study.

Keywords Accident Analysis; Crashworthiness; Euro NCAP; Injury Probability; Statistics

Crash testing is a way to determine whether best practice in occupant protection has been implemented in a new car. The European New Car Assessment Program (Euro NCAP) is a crash test program which was set up in 1996. From 1996 until spring 2000, 64 different car models had been tested and the results published. As part of this program cars are tested in both a frontal and side collision and, in 2000, a pole test was also introduced. Pedestrian protection is also considered in these tests. The present study considered only the front seat occupant protection aspect of the Euro NCAP scores (Hobbs et al., 1999). The test set-up and results can be studied on the Internet at http://www.euroncap.com.

The aim of the Euro NCAP crash test program is twofold. First, it provides objective information for the consumer. But it also promotes manufacturers who make an effort to improve their vehicles beyond the demands of legislation. Crash testing offers an early indication of the safety level of new cars. However, when cars have been on the market for some time, data obtained from real-life accidents provide important and more valid information about the real-life protection level of cars.

The Euro NCAP uses stars to indicate the safety level of vehicles. The star rating summarizes the combined protection level in front and side collisions. The star rating is produced using point scores from the front and side collisions: A maximum of 34 points can be achieved, 16 points from the frontal and 18 points from the side collision test. The intention of the scores is to give an indication of the extent to which best practice has been implemented in an individual car model, rather than predict the real-life crash outcome. Neither the test set-up nor the scoring system would be theoretically able to predict the outcome in all types of crashes. However, there should be a good correlation between higher scoring in the Euro NCAP tests and overall safety benefits in real-life accidents.

The aim of this study was to determine whether there is a correlation between successful application of best practice, as shown by the Euro NCAP front and side protection scores, and benefits in real-life impacts. A further aim was to apply new statistical techniques to generate injury risk functions to evaluate

Received 5 December 2000; accepted 14 March 2002.
Address correspondence to Anders Lie, Swedish National Road Administration, Röda Vägen 1, S-781 87 Borlänge, Sweden. E-mail: anders.lie@vv.se

the consistency of reduced injury risks for vehicles obtaining high Euro NCAP scores. While Euro NCAP considers both child occupant protection and pedestrian safety, neither of these aspects was considered in this study.

METHODS

The paired comparison statistical analysis method was employed in this study, where two car accidents were used to create relative risks (controlling for impact severity). The method was originally developed by Evans (1986), but has been developed further for car-to-car collisions by Hägg et al. (1992). The relative injury risk for the case vehicle is calculated by comparing the injury outcome in that vehicle with the injury outcome in the opponent vehicle. In car-to-car crashes, the mass proportions play a role as these influence the impact severity. This can be accounted for in the model. On this basis, the risk figures are sensitive (apart from the passive safety) only to systematic differences in seat belt use and accident type, which do not seem to be likely sources of error in this study. The method assumes that injuries in one car are independent from the injuries in the other car, given an impact severity.

In the paired comparison method, crash outcomes in two-car crashes are grouped into four groups (Figure 1 shows the layout); x_1 (injuries in both cars), x_2 (injuries in the case car but not in the opponent), x_3 (injuries in the opponent vehicle but not in the case vehicle), and x_4 (no one is injured in the crash, and usually no data is available). When calculating relative risks, x_4 does not add any important information and therefore is not used.

The risk relation between the two cars is calculated as the quotient between injuries in the case vehicle compared with the injuries in the opponent vehicle by using Eq. (1). The opponent vehicle is considered to be a sample of the whole vehicle population and therefore is the exposure measure, allowing comparisons across all case vehicles.

$$R_1 = (x_1 + x_2)/(x_1 + x_3) \tag{1}$$

If there is a weight difference between the case vehicle and the opponent vehicle, both vehicles will be exposed to an impact severity that differs compared to crashes where the opposing vehicles have the same weight (Figure 2). If the case vehicle is lighter than the average vehicle it will have a higher change of velocity compared to the average vehicle. At the same time, the

		Opponent vehicle	
		Injured	Not injured
Case car	Injured	x_1	x_2
	Not injured	x_3	x_4 *(unknown)*

Figure 1 Grouping of car-to-car cases into x_1, x_2 and x_3 sums.

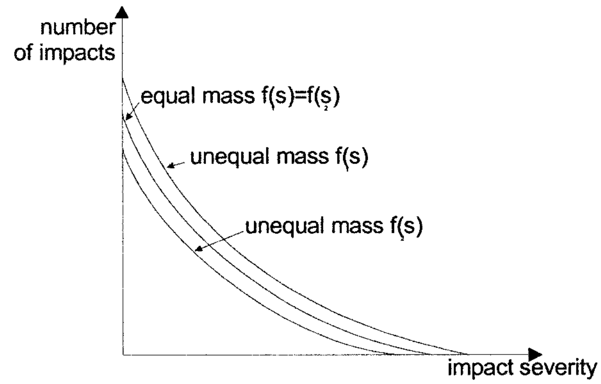

Figure 2 Change of velocity distribution depending on mass.

opponent vehicle will have a lower change of velocity. The mass therefore has a double effect on the relative risk. While it might be desirable to take into account the importance of weight for the case vehicle, the altered impact severity distribution must be compensated for relative to the opponent vehicle in order to allow comparisons with other case vehicles. The weight/change of velocity compensation factor was derived from an analysis of a defined set of vehicles that had been affected by varying mass in crashes. It was found that the risk of any injury, as well as severe and fatal injury, was increased or decreased by 7% for every 100 kg difference from the average weight. The compensation was performed using Eq. (2).

$$R_{\text{comp}} = R_1 \frac{1.07^{((M_{\text{case}} - M_{\text{avg}})/100)}}{1.07^{((M_{\text{opp}} - M_{\text{avg}})/100)}} \tag{2}$$

Crash testing into a fixed barrier is equivalent to a crash into a car of the same mass, while in real-life crashes weight is a factor that influences impact severity. In order to completely remove mass effects from the analysis, thereby allowing direct comparisons between barrier crash tests and real-life crashes, the mass of the case vehicles should be compensated for using the compensation factor (7% per 100 kg). This compensation was done using Eq. (2), but substituting the original R_1 value with the R_{comp} value of the first compensation. The standard deviation calculations were based on Gauss's approximation of variance for ratios.

The paired comparison method, as described above, calculates the average injury risk for a vehicle model. Based on the same data, information can be derived about the relative risk for a change of velocity (Krafft et al., 2000). The method used to derive risk functions uses the difference in mass between two cars in a car-to-car crash. The change of velocity for the individual vehicle depends on the relative speed and mass proportions [Eq. (3)]. This is based on the law of the conservation of momentum.

$$\Delta v = v_{\text{rel}}(M_2/(M_1 + M_2)) \tag{3}$$

By analyzing the risk in the case vehicle, when it collides with a range of opponent cars with varying known masses, risk curves can be derived. The derived curves are related to the average change of velocity and the average risk. The risk curves can cover a range of ± 15 to 20% from the average change of velocity for the injury severity studied. The risk functions show the elasticity in relative risks for relative changes of velocity. It is not possible to generate absolute figures for the risk functions without bringing in a key value (estimated or calculated). The key value can be either relative velocity or change of velocity. While this information is not recorded in police data, the relations can be relative only to each other. In practice, cars with different rating levels were analyzed by separating the opponent vehicles into mass categories, thereby calculating the elasticity to mass relations that can be translated to relative change of velocity. The method is described in detail in Krafft et al. (2000).

MATERIALS

This study was based on police reports from crashes in Sweden between January 1, 1994 and March 15, 2000. Only car-to-car crashes with known car makes and models were included. The police in the field classified the injuries according to United Nations Economic Commission for Europe (UNECE) definitions. Four injury levels were used: no injury, minor injury, severe injury (typically admitted to hospital), and fatal injury. Only injured drivers were studied. The data was analyzed in two groups: one containing only severe and fatal injuries; and the other containing fatal, severe, and minor injuries together. The data set contains information about vehicle make and model together with injury data for all crashes.

The vehicles were grouped by Euro NCAP star ratings (the corresponding point scores are available). The points from the front and side tests were used with star borders at 8, 16, 24, and 32 points. Until spring 2000, a maximum of four stars was possible. A fifth star can now be achieved if the point score is 32 points or more. As of May 2000, no car had yet achieved this result. Prior to publication, Euro NCAP recalculates the scores to protect level percentages.

Because of the limited numbers, all cars with the same star rating were grouped independently of their size group. The curb weight for every individual car was collected from the vehicle register. Only one car model has achieved a Euro NCAP one-star score, and that specific car model has not been sold in Sweden. All cars with Euro NCAP scores were used and scores from Euro NCAP phases 1 to 7a were used (i.e., all tests published before May 2000). Vehicles without Euro NCAP scores were used as a reference group. For this group, only 1994 year model cars and later were used. For the opponent vehicles in the pairs, all vehicles used had a curb weight between 700 kg and 2,500 kg. In total, 1,779 cases with severe or fatal injury outcome were studied along with 12,214 cases with at least a minor injury outcome. Of the vehicles with a known Euro NCAP score, 20 drivers were killed, 273 were severely injured, and 2,172 sustained a minor injury. In total, 15,901 car-to-car crashes were used for the mass compensation model. These vehicles were grouped in 100 kg groups.

RESULTS

Vehicles with different star ratings were found to have different injury risks in real-life crashes when severe and fatal injuries were considered. The distribution of severe and fatal injuries in the case and opponent vehicles is shown in Table I. The relative injury risk can be calculated using the figures in Table I (Eq. 1). The relative risk of sustaining a fatal or severe injury was calculated as well as the risk of sustaining minor through to fatal injuries. The average mass for the case and opponent vehicles was calculated. The Euro NCAP score for cars with a Euro NCAP rating was summed and average values were calculated (Table II).

Table I Distribution of injuries in case vehicle and opponent vehicle (original data for the calculations)

Severe and fatal injuries	All other vehicles		Minor, severe, and fatal injuries	All other vehicles	
Cars without Euro NCAP rating					
$n = 1{,}227$	Injured	Not injured	$n = 8{,}460$	Injured	Not injured
Injured	(x_1) 343	(x_2) 411	Injured	(x_1) 2,688	(x_2) 2,867
Not injured	(x_3) 473	(x_4) Unknown	Not injured	(x_3) 2,905	(x_4) Unknown
Euro NCAP two-star cars					
$n = 226$	Injured	Not injured	$n = 1{,}534$	Injured	Not injured
Injured	(x_1) 55	(x_2) 80	Injured	(x_1) 486	(x_2) 574
Not injured	(x_3) 91	(x_4) Unknown	Not injured	(x_3) 474	(x_4) Unknown
Euro NCAP three-star cars					
$n = 267$	Injured	Not injured	$n = 1{,}866$	Injured	Not injured
Injured	(x_1) 68	(x_2) 59	Injured	(x_1) 596	(x_2) 577
Not injured	(x_3) 140	(x_4) Unknown	Not injured	(x_3) 693	(x_4) Unknown
Euro NCAP four-star cars					
$n = 59$	Injured	Not injured	$n = 354$	Injured	Not injured
Injured	(x_1) 20	(x_2) 11	Injured	(x_1) 119	(x_2) 113
Not injured	(x_3) 28	(x_4) Unknown	Not injured	(x_3) 122	(x_4) Unknown

Abbreviation: Euro NCAP, European New Car Assessment Program.

Table II Relative risk values, average mass, and average Euro NCAP point values

Stars	R_1 Severe and fatal injuries	R_1 All injuries	Avg. mass case (kg)	Avg. mass opponent (kg)	Avg. Euro NCAP points
No class	0.92	0.99	1,332	1,287	
2	0.92	1.10	1,260	1,288	13.09
3	0.61	0.91	1,450	1,297	21.06
4	0.65	0.96	1,362	1,304	25.98

Abbreviation: Euro NCAP, European New Car Assessment Program.

When the effects of mass on impact severity distribution were compensated for (Eq. 2), the risk values changed slightly (Table III). This compensation was performed in two steps: first, where there were vehicle mass differences in real-life crashes, compensation was made to adjust the safety level. In the second step, the compensation eliminated the mass effect. The values derived after the second compensation are expected to best reflect the outcome in the Euro NCAP tests since the laboratory test mass is neutral and simulates the case when the case vehicle collides with a vehicle of the same mass, at least in the frontal impact crash.

Figure 3 shows the relative risk after the calculation of the mass compensation. Since all mass effects have been removed, the results are directly comparable with the crash test results. The Euro NCAP results were plotted as the average score. It can be seen that for the two- and three-star rated cars the average score was higher than the median score within the star range (the median being 12 respectively 20 points). The four-star group had an average score under the median score (the median being 28 points) for the star span.

When a straight line is fitted to Figure 3, values for the median score in a star band can be calculated. The line shows a 12% risk reduction per star for severe or fatal injuries. No difference was seen for the minor injuries. This procedure shows the estimated risk level for severe and fatal injuries for cars with 4, 12, 20, and 28 points in the Euro NCAP scoring system (Table IV).

RISK FUNCTIONS

In Figures 4 and 5, relative risk functions (relative injury risk vs. relative change of velocity) are shown for vehicles with either no rating or two-star ratings versus vehicles with a three-

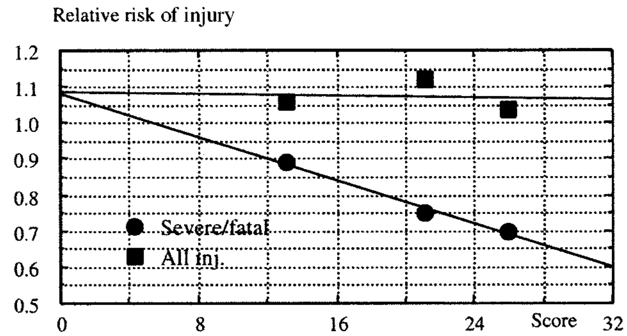

Figure 3 Relative injury risk after mass compensation to crash test conditions.

or four-star rating. The limited numbers of cases available and used made this grouping necessary. It can be seen that there is a consistency in the relation between the two groups of vehicles, and that the difference is large for both lower-end and higher-end crashes although the difference seems to be larger for the more severe impacts. The analysis allows studies of only a fairly small range of impact severity. However, it can be seen that over this range, the three- and four-star rated vehicles would have to be exposed to a change of velocity that is approximately 12% higher than the no rating and two-star rated cars to generate the same risk for serious or fatal injury. It also can be seen in Figure 5 that there is no clear difference between these groups when minor injuries are studied. Please note that the scales in Figures 4 and 5 cannot be compared.

DISCUSSION

Euro NCAP is an initiative that aims to drive vehicle safety beyond current regulation requirements by offering the consumer more information about the safety features of cars. Euro NCAP does not and cannot predict real-life crash outcome on a car-by-car basis. It should not even theoretically be able to do that in the current form with a star ranking system with no reflection on representative weights on different aspects. However, it is still important to evaluate whether the aim of promoting vehicles with higher levels of safety features is valid. The most effective way to do this is to compare real-life crash outcomes with the ratings and scores provided by Euro NCAP. While this

Table III Relative risk compensated for mass influence to reflect real-life and crash test conditions

Stars (Euro NCAP)	R_1 Severe and fatal injuries (SD)	R_1 Severe and fatal, mass compensated to real life[a]	R_1 Severe and fatal, mass compensated to crash test[b]	R_1 All injuries (SD)	R_1 All, mass compensated to real life[a]	R_1 All, mass compensated to crash test[b]
No class	0.92 (0.01)	0.95	0.98	0.99 (0.01)	1.02	1.06
2	0.92 (0.03)	0.91	0.89	1.10 (0.02)	1.08	1.06
3	0.61 (0.03)	0.68	0.75	0.91 (0.02)	1.01	1.12
4	0.65 (0.03)	0.67	0.70	0.96 (0.03)	1.00	1.04

Note. Standard deviation (SD) within parentheses. The magnitude of standard deviation is the same through the compensations.
Abbreviation: Euro NCAP, European New Car Assessment Program.
[a]Compensation has been made only for the opponent vehicle with 7% per 100 kg.
[b]Compensation has been made also to the target vehicle on the same level with 7% per 100 kg.

Table IV Predicted relative risk for the center of the Euro NCAP star bands, compensated for mass influence to crash test conditions

Center of star	R_1 Severe, mass compensated to crash test conditions
1 (extrapolated)	(1.02)
2	0.90
3	0.78
4	0.66

Abbreviation: Euro NCAP, European New Car Assessment Program.

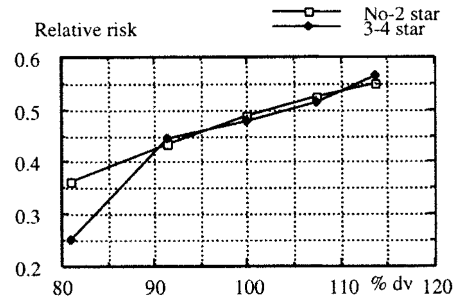

Figure 5 Risk functions for all injuries to relative change of velocity.

study shows a very good overall relationship, it should not be seen as proof that there is a predictive value in the Euro NCAP system, especially not for individual car model scores. Other reasons could explain this general relationship. For example, manufacturers who develop vehicles with high safety standards generally do well in Euro NCAP. However, this does not mean that a vehicle that was designed entirely for good Euro NCAP results will also perform well in real-life crashes. On the other hand, a vehicle that performs well in real-life may indeed also obtain a high Euro NCAP score. Similar results have been seen in other parts of the world (Kahane et al., 1994; Newstead & Cameron, 1998; O'Neill et al., 1994).

While Euro NCAP is a process that should encourage manufacturers to incorporate best safety practice, and may further develop best practice, it is important that the real-life crash outcome of vehicles is constantly monitored. Such monitoring is important to find indications of suboptimization, or negative consequences of vehicles being designed to achieve a good Euro NCAP score. Analysis of real-life performance data could help pinpoint car models that are built only for good results in crash tests.

There has been concern that the test speeds used in Euro NCAP crash tests might lead to suboptimization, possibly leading to better vehicle performance in high-severity crashes and worse performance in low-severity impacts. The analysis of risk functions performed in this study does not support this concern; however, the issue should be monitored. It also would be beneficial to monitor any sign of reduced vehicle compatibility. Compatibility has been raised as a possible concern; however, earlier studies have suggested that this is not a necessary consequence (Lie et al., 1996).

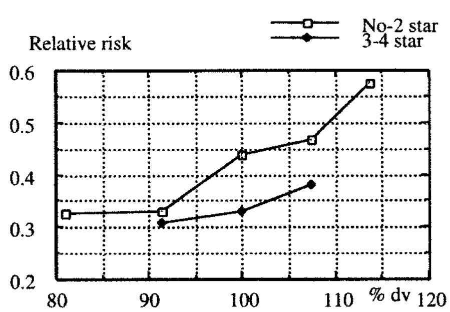

Figure 4 Risk functions for severe and fatal injuries to relative change of velocity.

Even if all crashes over several years were included, there are still relatively few crashes available in Sweden for this kind of study. A more detailed study, using the same methodology, could be performed with a larger data set. If all European countries could merge their data, more reliable and precise results could be achieved. Risk levels for every size group and for individual car models could then be calculated.

In the Swedish police crash data, no information is included about the specific injuries or the point(s) of impact. If such data were available, an indication of the front and side protection levels could be included and individual types of injuries could be studied. Further research could look for relationships between risk levels within the different vehicle size groups used by Euro NCAP. Also the risk levels for individual car models could be studied if larger data sets were available. This could help in the comparison between groups of vehicles of different sizes. Further verification could also be achieved if some old cars with known real-life performance were tested.

No major differences could be seen concerning minor injuries. This confirms earlier research, and suggests that the focus of the Euro NCAP and more generally in vehicle development, is toward a reduction of more severe injuries. This is further supported by the results of real-life crashes. It is of concern, however, that the results did not show any benefit to minor injuries. While these might be seen as being less important to reduce, they still contain some injury types that generate long-term health losses. For example, neck injuries in rear-end and frontal collisions have traditionally been defined as minor injuries; however, evidence suggests that they do contribute to a significant loss of health (Krafft, 1998). It is important that the Euro NCAP in the future focus on these kinds of injuries as well.

It is clear that the reductions achieved in serious and fatal injuries are substantial. It is to some extent surprising that it is possible to discriminate between cars built at the same time, and with differences that are at a level where they can influence safety. The magnitudes of the safety differences are at a level that they become one of the major instruments for the future of traffic safety. While there was no difference between cars that were ranked with two stars and older vehicles, four-star cars seem to reduce the risk of serious and fatal injuries by more than 30%.

When applying risk functions to the statistical data, it was shown that in order generate the same risk for serious and fatal injuries for the low- and high-rated vehicles, the change of velocity would have to be increased by approximately 12% for the three- and four-star group. This could be seen as an indication as of the level of higher test speeds now used. The risk functions did not show that higher-performing vehicles produced more injuries in lower impact severity crashes, although the method used did not allow consideration of very low severity impacts.

The importance of weight should not be underestimated, and while this factor is not taken into account in crash tests into fixed barriers, 100 kg more weight in a car-to-car impact will generate a 7% lower risk of injury. In single vehicle crashes, which account for a high proportion of crashes, the mass should not have any significant influence on safety.

CONCLUSIONS

There was an overall correlation between the Euro NCAP scores and risk of serious and fatal injury. The results indicate a 12% per star risk reduction for severe and fatal injuries. No overall relationship was found between Euro NCAP scores and minor injury crashes. Highly rated vehicles, as a group, had a lower risk of serious and fatal injury across 90–110% of average impact severity, indicating that in crashes of such severity there have not been any drawbacks of the high test speed in Euro NCAP. Overall, highly rated vehicles produce approximately 30% less fatal and serious injuries compared with low-rated vehicles.

REFERENCES

Evans, L. (1986) Double Paired Comparison—A New Method to Determine How Occupant Characteristics Affect Fatality Risk in Traffic Crashes, *Accid. Anal. Prev.*, Vol. 18, pp. 217–227.

Hägg, A., von Koch, M., Kullgren, A., Lie, A., Nygren, Å., Tingvall, C. (1992) "Folksam Car Model Safety Rating 1991–92," Folksam Research, Stockholm.

Hobbs, C., Gloyns, P., Rattenbury, S. (1999) "Euro NCAP, Assessment Protocol and Biomechanical Limits, Version 2," European New Car Assessment Program, Brussels.

Kahane, C. J., Hackney, R. J., Berkowitz, A. M. (1994) Correlation of Vehicle Performance in the New Car Assessment Program with Fatality Risk in Actual Head on Collisions, *Proc. ESV Experimental Safety Vehicle*, Munich 1994. Paper no. 94-S8-O-11, Vol. 2, pp. 1388–1404.

Krafft, M. (1998) Non-Fatal Injuries to Car Occupants: Injury Assessment and Analysis of Impacts Causing Short- and Long-Term Consequences with Special Reference to Neck Injuries, Ph.D. dissertation, Karolinska Institute, Stockholm.

Krafft, M., Kullgren, A., Les, M., Lie, A., Tingvall, C. (2000) Injury as a Function of Change of Velocity: An Alternative Method to Derive Risk Functions, *Proc. IMechE, Vehicle Safety*, pp. 263–273.

Lie, A., Tingvall, C., Larsson, P. (1996) The Crash Safety of New Car Models—A Comparative Accident Study of New Versus Old Car Models, *Proc. ESV Experimental Safety Vehicle*, Melbourne 1996. Paper no. 96-S9-O-13, Vol. 2, pp. 1441–1443.

Newstead, S., Cameron, M. (1998) "Updated Correlation of Results from the Australian New Car Assessment Program with Real Crash Data from 1987 to 1996," Monash University Accident Research Centre (MUARC) Report. Melbourne, Australia.

O'Neill, B., Lund, A. K., Zuby, D. S., Preuss, C. A. (1994) Offset Frontal Impacts—A Comparison of Real-World Crashes with Laboratory Tests, *Proc. ESV Experimental Safety Vehicle*, Munich 1994. Paper no. 94-S4-O-19, Vol. 1, pp. 649–670.