HOW DOES EURO NCAP RESULTS CORRELATE TO REAL LIFE INJURY RISKS - A PAIRED COMPARISON STUDY OF CAR-TO-CAR CRASHES

Anders Lie – Swedish National Road Administration and Karolinska Institutet, Sweden
Claes Tingvall – Monash University Accident Research Centre, Australia

ABSTRACT

Euro NCAP is a resource for consumers regarding vehicle crash safety. The program also promotes safety developments, and credits car manufacturers focussing on safety. This study, based on real life car to car crashes, shows that the overall indication of the safety level, provided by the crash testing, is a valid prediction, at least when looking at the star rating and severe to fatal injuries. For minor injuries no significant injury risk differences are seen. The cars with three or four stars are approximately 30% safer, compared to two star cars or cars without an Euro NCAP score, in car to car collisions. The good general correlation between injury risk, and Euro NCAP scores is not necessarily similarly good for individual car models. Pedestrian safety and child occupant protection was not studied.

Key words; Accident analysis, crashworthiness, injury probability, statistics, Euro NCAP

CRASH TESTING OF VEHICLES is a way to determine if best practice in terms of occupant protection has been implemented for a new car. Euro NCAP is a crash test program, which was set up in 1996. Since that time and up until springtime 2000, 64 different car models have been tested and the results have been published. The cars are tested in a frontal collision and in a side collision. The possibility of adding a pole test has been introduced 2000. A pedestrian protection test is also included. This study only covers the front seat occupant protection part of the Euro NCAP scores (Hobbs et al.1999). The test set-up and results can be studied on the Internet site www.euroncap.com.

The aim of the Euro NCAP crash test program is two fold. There is a need for objective consumer information, but there is also a need to promote industry when an effort is made to improve their vehicles beyond the demands of legislation. Crash testing is a way to get an early indication of the safety level of new cars. When cars have been on the market for some time, ratings obtained from real life accidents give important and more valid information about the real life protection level of cars.

Euro NCAP uses stars to indicate the safety level of a vehicle. A combined star rating shows the protection level in the front collision and side collision together. The star scoring is based on point scores for the front and side. Maximum 34 points can be achieved by adding 16 front and 18 side points. The intention of the scores is to give an indication to what extent best practice or benchmarking has been applied to an individual car model, and not to predict the real-life outcome. Neither the test set-up nor the scoring system would be theoretically able to predict the outcome in all types of crashes. On the other hand, there should be a good correlation between promoting higher scoring in the Euro NCAP and overall safety benefits in road accidents.

The aim of this paper is to find to what extent there is an overall correlation between successful application of best practice as shown by Euro NCAP front and side protection scores, and benefits in real life impacts. The aim was also to apply new statistical techniques to generate injury risk functions to evaluate the consistency of reduced injury risks for high scored vehicles. While Euro NCAP, as well as child occupant protection address pedestrian safety, neither of these aspects were included in this study.

METHODS

The basis for the statistical analysis method is the paired comparison technique, where two car accidents are used to create relative risks. The method was initially developed by Evans, but has been developed further for car to car collisions by Hägg et al. (Evans 1986, Hägg et al. 1992). The paired comparison method controls for impact severity. The relative injury risk for the case car model is calculated by comparing the injury outcome in the case vehicle with the injury outcome in the opponent vehicles, it crash with. In car to car crashes the mass proportions play role, as they will alter the impact severity distribution. This can be taken into account in the model. On this basis the risk figures are only sensitive (apart from the passive safety) 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 a certain impact severity.

In the paired comparison method, crash outcomes in two-car crashes are grouped in four groups (Fig. 1) x_1 (injuries in both cars), x_2 (injuries in the case car but not in the opponent) and x_3 (injuries in the opponent vehicle but not in the case vehicle). If no one is injured in the crash, x_4, usually no data is available. To calculate relative risks, x_4 is not used and will not add any important extra information.

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

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

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

$$R_1=(x_1+x_2)/(x_1+x_3) \qquad \text{Eq. 1.}$$

If there is a weight difference between case vehicle and opponent vehicle, both vehicle types will be exposed to impact severity that differs from when the two groups of vehicles have the same weight (Fig. 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 will have a lower change of velocity. The mass has therefore a double effect on the relative risk. While it might be desired to take into account the importance of weight for the case vehicle, the altered impact severity distribution must be compensated for relating to the opponent vehicle in order to allow comparisons with other case vehicles. Analysing how a defined set of vehicles was affected by varying mass relations derived the weight/change of velocity compensation. 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 equation 2.

$$R_{comp} = R_1 \frac{1{,}07^{((M_{case}-M_{avg})/100)}}{1{,}07^{((M_{opp}-M_{avg})/100)}} \qquad \text{Eq. 2.}$$

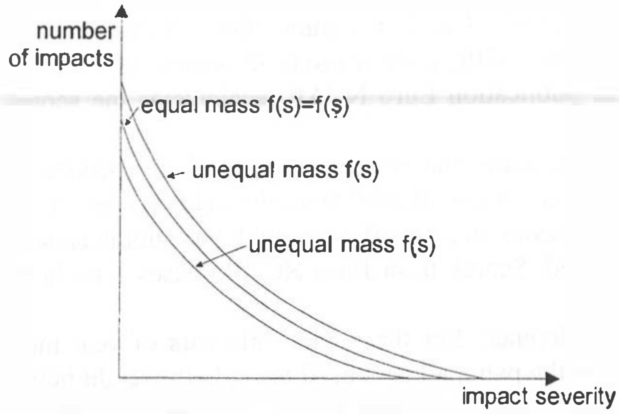

Figure 2. Change of velocity distribution depending on mass.

Crash testing into a fixed barrier is equivalent to a crash into a car of the same mass, while the real life outcome integrates weight as a factor influencing impact severity. In order to completely remove mass effects from the analysis, thereby allowing direct comparisons with barrier crash tests, it is possible to compensate for mass for the case vehicles as well. The same mass compensation factor was used (7% per 100kg). Also this compensation was done using equation 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' approximation of variance for ratios.

The paired comparison method, as described above, describes the average injury risk for a car model. Based on the same data, information can be derived about the relative risk level over a change of velocity range (Krafft et al. 2000). The method 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.

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

By analysing the risk in the car model under study when it collides with a range of opponent cars with varying known masses, risk curves can be derived. The derived curves are only 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 either be relative velocity or change of velocity. While they are not known in police records, the relations can only be relative to each other. In practice, cars with different rating levels were analysed by dividing 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).

MATERIAL

The study is based on police reports from crashes in Sweden between 1994-01-01 to 2000-03-15. Only car to car crashes with known car makes and models are included. The police in the field have classified the injuries in correspondence with the ECE definitions. Four injury levels are used; no injury; minor injury; severe injury; and fatal injury. The severe injuries should typically lead to hospital admittance. Only injured drivers are studied. The data was analysed in two groups: one containing only severe and fatal injuries; and one 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 cars are grouped by Euro NCAP star ratings. The Euro NCAP point score for the rated cars are available. The sum of points, from front and side tests is used. Euro NCAP uses 0-34 points to do the star rating. The star borders are at 8, 16, 24 and 32 points. Until springtime 2000 a maximum of four stars has been possible. Now a fifth star can be achieved if the point score is 32 points or more. No car has achieved that result yet (May 2000). Before publication Euro NCAP recalculates the scores to protection level percentages.

Because of the limited material, all cars with the same star rating are grouped independently of their size group. The kerb weight for every individual car is collected from the vehicle register. Only one car model has achieved an Euro NCAP one star score, that specific car model has not been sold in Sweden. All cars with Euro NCAP scores were used. Scores from Euro NCAP phases 1 to 7a were used, that is all tests published before May 2000.

Cars without Euro NCAP scores are used as reference. For these cars only cars of year models from 1994 and later were used. As opponent cars in the pairs, all cars used had a kerb weight between 700kg and 2500kg. In total 1779 cases with severe or fatal outcome were studied along with 12214 cases with at least minor injury outcome. In cars with known Euro NCAP score, 20 drivers were killed, 273 were severely injured and 2172 sustained a minor injury.

In all 15901 car to car crashes were used for the mass compensation model. These cars were grouped in 100kg groups.

RESULTS

The cars with different star ratings have different injury risk in real life crashes when severe and fatal injuries are studied. The distribution of severe and fatal injuries, in the case vehicle and the opponent vehicle is shown below.

Cars without Euro NCAP rating

Severe and fatal injuries	All other vehicles		Minor, severe and fatal inj.	All other vehicles	
n=1227	Injured	Not injured	n=8460	Injured	Not injured
Injured	(x_1) 343	(x_2) 411	Injured	(x_1) 2688	(x_2) 2867
Not injured	(x_3) 473	(x_4) *unknown*	Not injured	(x_3) 2905	(x_4) *unknown*

Euro NCAP 2 star cars

Severe and fatal injuries	All other vehicles		Minor, severe and fatal inj.	All other vehicles	
n=226	Injured	Not injured	n=1534	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 3 star cars

Severe and fatal injuries	All other vehicles		Minor, severe and fatal inj.	All other vehicles	
n=267	Injured	Not injured	n=1866	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 4 star cars

Severe and fatal injuries	All other vehicles		Minor, severe and fatal inj.	All other vehicles	
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*

Table 1. Distribution of injuries in case vehicle and opponent vehicle.

From the values above the relative injury risk can be calculated using Eq. 1. The relative risk to sustain a fatal or severe injury is calculated as well as the risk to sustain minor to fatal injuries. During the calculation, average mass for the case cars as well as the opponent cars is calculated. The

Euro NCAP score for cars with Euro NCAP rating is summed and average values are calculated (Table 2).

Stars	R_l Severe and fatal injuries	R_l All injuries	Avg. mass case (kg)	Avg. mass opponent (kg)	Avg. Euro NCAP points
No class	0,92	0,99	1332	1287	
2	0,92	1,10	1260	1288	13,09
3	0,61	0,91	1450	1297	21,06
4	0,65	0,96	1362	1304	25,98

Table 2. Relative risk values, average mass and average Euro NCAP point values.

If compensation for mass effects on impact severity distribution is included (Eq. 2), the risk values change slightly, table 3. The compensation is performed in two steps: first compensation is performed to adjust to the safety level in real life traffic where the mass differences for the case vehicles are accounted for. In the second step the compensation eliminates the mass effect. The values derived after the second compensation should best reflect the outcome in the Euro NCAP tests since the laboratory test mass is neutral and simulates the case when the case car collides with a car of the same mass, at least in the frontal impact.

Stars	R_l Severe and fatal injuries (s.d.)	R_l Severe and fatal Mass compens. to real life[1]	R_l Severe and fatal Mass compens. to crash test[2]	R_l All. Injuries (s.d.)	R_l All. Mass compens. to real life[1]	R_l All. Mass compens. to crash test[2]
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

Table 3. Relative risk compensated for mass influence, to reflect real life and crash test conditions. Standard deviation (s.d.) within parentheses.

[1] *Compensation has only been made for the opponent vehicle with 7% per 100kg.*

[2] *Compensation has been made also to the target vehicle on the same level with 7% per 100kg.*

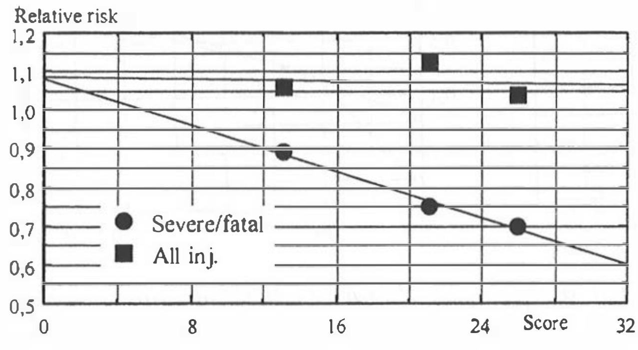

Figure 3. Relative injury risk after mass compensation to crash test conditions

Figure 3 shows the relative risk after mass compensation. Since all mass effects are removed the results are comparable with the crash test results. The crash test simulates a crash with a vehicle of the same mass. The Euro NCAP results are plotted as the average score. It can be seen that for the two and three star cars the average score is higher than the median score value within the star range (12 respectively 20 points). The four star group has an average score under the median score (28 points) for the star span.

If a straight line is fitted in figure 3, values for the median score in a star band can be calculated. This operation shows the estimated risk level for severe and fatal injuries for cars having 4, 12, 20 and 28 points in the Euro NCAP scoring system (Table 4).

Centre of star	R_I Severe. Mass compens. to crash test
1 (extrapolated)	(1,02)
2	0,90
3	0,78
4	0,66

Table 4. Predicted relative risk for the centre of the star band, compensated for mass influence to crash test conditions

RISK FUNCTIONS

In figures 4 and 5, relative risk functions (relative injury risk versus relative change of velocity) for cars without a rating (0 stars) or rated 2 stars, versus cars rated 3 or 4 stars are shown. The limited material 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 only allows studies of a fairly small range of impact severity. However, it can be seen that over this range, the 3 and 4 star ranked vehicles would have to be exposed to a change of velocity that is approximately 12% higher than the 0 and 2 star ranked cars to generate the same risk for serious to fatal injuries. It can also be seen in figure 5 that there is no difference when minor injuries are included, although these on the general level are more or less as sensitive to impact severity as the serious and fatal injuries.
Please note that the scales in figures 4 and 5 can not be compared.

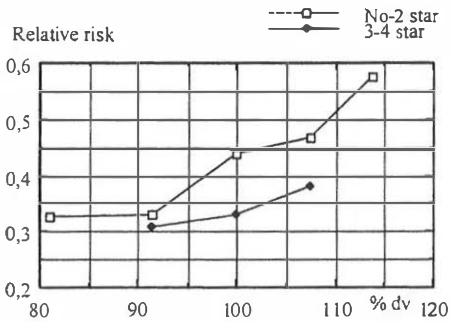

Figure 4. Risk functions for severe and fatal injuries to relative change of velocity

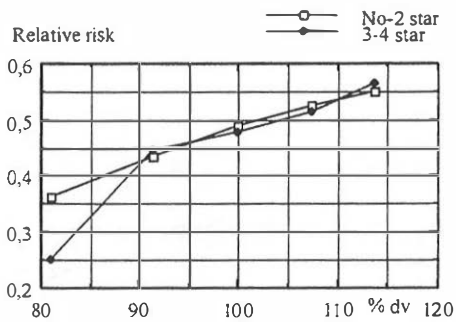

Figure 5. Risk functions for all injuries to relative change of velocity

DISCUSSION

Euro NCAP is an initiative to drive vehicle safety beyond current regulation, by offering the market more extensive information about adoption of best practice. Euro NCAP has never been intended to predict real life 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 general aim of promoting vehicles with higher safety standards is fulfilled. The most natural way to do this is to compare real life outcomes with the ratings or scores in Euro NCAP. While this study shows a very good overall relationship, it should not be seen as proof that there is a predictive value in Euro NCAP, especially not for individual car model scores. There might be several reasons for this general relationship, the most likely being that car manufacturers developing cars with high safety standards also do well in Euro NCAP. That does not mean that a vehicle that was designed entirely for good Euro NCAP results will perform well in real life crashes. It is, however, clear that a car that performs well in real life also can be highly scored in Euro NCAP, something that is becoming increasingly important. Similar results have been seen in other parts of the world (Newstead and Cameron 1998, Kahane et al 1994, O'Neill et al 1994).

While Euro NCAP is a process that should lead to car manufacturers aiming for best practice, and does develop and drive further development of best practice, it is of importance that the real life outcome is constantly monitored. Monitoring is of importance also in finding any indications of sub-optimisation or negative consequences of designing cars to create good NCAP scoring. A co-operation around real life follow-up, could help pinpoint car models that are built only for good results in crash test.

There has been concern that the test speeds in Euro NCAP might lead to sub-optimisation, possibly leading to better vehicle performance in high severity crashes and worse in low severity impacts. The analysis of risk functions in this study does not suggest that this is the case at this moment. However, it is still important to monitor this issue. It would also be beneficial to study any sign of reduced compatibility. Compatibility has been raised as a possible concern, but earlier studies have shown that this is not a necessary consequence (Lie et al 1996).

Even if all crashes over several years were included, relatively few crashes are available for this kind of study. When looking at severe and fatal crashes in a country like Sweden, this is evident. It is though important to know that the crashes used in this study, is a sample from a much larger number of crashes. 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 could then be calculated.

In the Swedish police material no information is given about the specific injury or the point of impact. If data were available some indication of the front and side protection levels could be added and individual types of injuries could be studied.

Further research could look for relations between risk levels within the different size groups used by Euro NCAP. Also the risk levels for individual car models can be studied if large data sets are available. This could help in the comparison between size groups. It would also be interesting to test some old cars with known real life performance. By doing this further verification could be achieved.

No major differences could be seen concerning minor injuries. This confirms earlier research, and suggests that the focus in Euro NCAP and more generally in car development, towards reduction of more severe injuries, can be seen also in real life crashes. There is, however, a concern that the results did not show any benefit to minor injuries. While these might be seen as less important to reduce, they still contain some injuries that generate long term health losses. This is especially relevant to neck injuries in rear end and frontal collisions, that traditionally are defined as minor injuries. However, they contribute to the total loss of health (Krafft 1998). It is important that the Euro NCAP process in the future focuses 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 2 stars and older vehicles, four star cars seems 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 3 and 4 star group. This could be seen as an indication as to the level, which cars have been designed to match as a result of the higher test speed than has been used before. The risk functions did not show that higher performing vehicles produced more injuries in lower impact severity segments, although the method used did not allow for looking at 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, 100kg 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 a strong and consistent overall correlation between Euro NCAP scoring and risk of serious and fatal injury
- No overall relationship between Euro NCAP scoring and minor injury could be detected
- High ranked 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, high ranked vehicles produce approximately 30% less fatal and serious injuries compared to low-ranked vehicles.

REFERENCES

Evans L. Double paired comparison – a new method to determine how occupant characteristics affect fatality risk in traffic crashes. Accident Analysis and Prevention, Vol. 18 No. 3, 1986, pp 217-227.

Hobbs C, Gloyns P, Rattenbury S. Euro NCAP, Assessment protocol and biomechanical limits. Version 2. 1999. Euro NCAP, Brussels 1999.

Hägg A, v Koch M, Kullgren A, Lie A, Nygren Å, Tingvall C, Folksam Car Model Safety Rating 1991-92, Folksam Research S-106 60, Stockholm, 1992.

Kahane C J, Hackney R J, Berkowitz. Correlation of vehicle performance in the new car assessment program with fatality risk in actual head on collisions. In proceedings ESV Experimental safety vehicle, Munich 1994. Paper no. 94-S8-O-11, Vol. 2, pp 1388-1404.

Krafft M, Kullgren A, Les M, Lie A and Tingvall C. Injury as a function of change of velocity, an alternative method to derive risk functions. In proceeding from IMechE, Vehicle Safety 2000. London 2000, pp 263-273.

Krafft M. Non-fatal injuries to car occupants. Injury assessment and analysis of impacts causing short- and long-term consequences with special reference to neck injuries. Thesis. Stockholm 1998

Lie A, Tingvall C, Larsson P. The crash safety of new car models – a comparative accident study of new versus old car models. In proceedings ESV Experimental safety vehicle, Melbourne 1996. Paper no. 96-S9-O-13, Vol. 2, pp 1441-1443.

Newstead S, Cameron M. 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 1998.

O'Neill B, Lund A K, Zuby D S, Preuss C A. Offset frontal impacts – a comparison of real-world crashes with laboratory tests. In proceedings ESV Experimental safety vehicle, Munich 1994. Paper no. 94-S4-O-19, Vol. 1, pp 649-670.